

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2003, 15:41:35 ; Search time 8038 Seconds
(without alignments)
11527.781 Million cell updates/sec

Title: US-09-928-175-1

Perfect score: 2265
Sequence: 1 atgattcgtttcttcgtttc.....taatgaaccagtttcctag 2265

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pac.*
7: gb_pn.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pac.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rnd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_by.*
39: em_hugo_hum.*
40: em_hugo_mus.*
41: em_hugo_other.*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2265	100.0	2265	6	AX385027	AX385027 Sequence
2	2265	100.0	2838	9	AF403384	AF403384 Homo sapi
3	2261.8	99.9	2436	9	AF453828	AF453828 Homo sapi
4	2258.8	99.7	2262	6	BD141180	BD141180 Novel G p
5	2214	97.7	2214	6	AX451562	AX451562 Sequence
6	2111	93.2	2193	6	AX385032	AX385032 Sequence
7	2104.8	92.9	2190	6	BD141181	BD141181 Novel G p
8	2060	90.9	2142	6	AX451566	AX451566 Sequence
9	2002.2	88.4	2181	6	AX709190	AX709190 Sequence
10	1803	79.6	2049	6	AX385037	AX385037 Sequence
11	1579.8	69.7	2539	10	AF346501	AF346501 Mus muscu
12	1579	69.7	2214	6	AX385045	AX385045 Sequence
13	1543.4	68.1	1545	6	BD141175	BD141175 Novel G p
14	1389.4	61.3	1473	6	BD141176	BD141176 Novel G p
15	1384.8	61.1	1830	6	BD141172	BD141172 Novel G p
16	1149	50.7	1149	6	AX385030	AX385030 Sequence
17	1066.4	47.1	1068	6	AX148176	AX148176 Sequence
18	1065	47.0	1065	6	BD141178	BD141178 Novel G p
19	995	43.9	1077	6	AX385035	AX385035 Sequence
20	991	43.8	1101	6	AX385042	AX385042 Sequence
21	825.8	36.5	1140	6	AX385048	AX385048 Sequence
22	735	32.5	1274	6	AX349337	AX349337 Sequence
23	735	32.5	2274	6	AF190500	AF190500 Homo sapi
24	735	32.5	2467	6	BD135245	BD135245 Novel mam
25	687	30.3	933	6	AX385040	AX385040 Sequence
26	686.2	30.3	3584	6	BD135246	BD135246 Novel mam
27	566.6	25.0	653	4	AY196483	AY196483 Equus cab
28	547	24.2	575	6	AX385204	AX385204 Sequence
29	482.2	21.3	1191	6	AX128507	AX128507 Sequence
30	413	18.2	44585	6	AX646897	AX646897 Sequence
31	413	18.2	170522	9	AL138708	AL138708 Human DNA
32	413	18.2	176565	2	AL136106	AL136106 Homo sapi
33	370	16.3	592	6	AX320420	AX320420 Sequence
34	321	14.2	321	6	AX384516	AX384516 Sequence
35	280.2	12.4	180673	10	AC068627	AC068627 Mus muscu
36	280.2	12.4	211251	2	AC121029	AC121029 Rattus no
37	280.2	12.4	213462	10	AC077689	AC077689 Mus Muscu
38	280.2	12.4	239190	2	AC098990	AC098990 Rattus no
39	232.4	10.3	576	6	AX320477	AX320477 Sequence
40	221	9.8	1018	6	AX147820	AX147820 Sequence
41	221	9.8	1018	6	AX521869	AX521869 Sequence
42	201.6	8.9	3756	3	LSGPR	223104 L.stagnalis
43	200	8.8	156490	10	AL603709	AL603709 Mouse DNA
44	200	8.8	174705	2	AC101958	AC101958 Mus muscu
45	184.4	8.1	187431	9	AC107219	AC107219 Homo sapi

ALIGNMENTS

RESULT 1
AX385027
LOCUS
DEFINITION
Sequence 1 from Patent WO0214489.
AX385027
VERSION
AX385027.1 GI:19578152
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
TITLE
Paszy, C.J., Gong, J., Daugherty, B. and Rogers, N.
Leucine-rich repeat-containing g-protein coupled receptor-8
molecules and uses thereof

JOURNAL Patent: WO 0214489-A 1 21-FEB-2002;
Amgen, Inc. (US)
FEATURES Location/Qualifiers
source 1..2265
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..2265
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28586.1"
/db_xref="GI:19578153"
/translation="MIVPLVFKHLFSLRLTWFLLHFLVILNVDFALTOGSMITPS
COKGFPCGNLTKCLPRAFHCDGKDCGNDCEBNCDSGMAITFGTVHGNANVAL
TOBFLKQYPOCCDCKELETCEVNDLKSVPIMSNVYLLSKKHLSIDPKYIKY
TKLKIPLQNHICIRHISRAKAFGLCNLIOLVHNHCITTLRPGIKDLOHQLTWLID
NPITISORLFTGLNSLPFLSMVNNYELALPKOMCAOMPOLNMDLBNRIRKYLNST
FLSCSLTVLPFLRNQIGFPEKPESSIKATGEDELSSNITTELSPHLFKDLKLOKLT
NLSNPLMYLHKNOPESSIKOQSDLERIEIPNINTRMFOPMKLSHYFKNRPSYCY
AMSTIKILCCADCLNGVLPFVGIPIKRGQYOKYALIMESVOCRLMGFLAMSTEV
SVLLITVLTLEKFLVIVPFSNIRPKGKTSVILICIMAGFLIAVLPKMKDFGNF
YKNGVCFPLVYDQTEDEIGSKYSLGIEIVNLAFILIVSYITMCSIQKIALQTT
EVKNCQFEVAVANRPFETIVSDAICWTPVAVKILSIFRVEIDPTMTSWIVPFLPV
NSALNPILYLTLPFKDKLKQDLHKIQRKSIKIKKSLSTSIWIEDSSSLKGLYL
NKITLGSIMKPVIS"

sig peptide 1..108
BASE COUNT 650 a 466 c 423 g 726 t
ORIGIN

Query Match 100.0%; Score 2265; DB 6; Length 2265;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGATGTTTTCGTTTAAACATCTCTCAAGCTCAATGATTCAATGTTCTTT 60
1 ATGATGTTTTCGTTTAAACATCTCTCAAGCTCAATGATTCAATGTTCTTT 60
61 CTACTTCATTTCACTGTTGATCATATGTCAGAAATTTTCACTCAAGTAGCATG 120
61 CTACTTCATTTCACTGTTGATCATATGTCAGAAATTTTCACTCAAGTAGCATG 120
121 ATCACTCCCTCAGGCCAAGATATTTCCCTGCGGAAATCTTACCAAGTCTTACC 180
121 ATCACTCCCTCAGGCCAAGATATTTCCCTGCGGAAATCTTACCAAGTCTTACC 180
181 CGAGCTTTTCACTGTAATGCAAGATGCTGCGGAAACGGGCGGACGAAGAACTGT 240
181 CGAGCTTTTCACTGTAATGCAAGATGCTGCGGAAACGGGCGGACGAAGAACTGT 240
241 GGTGACACTAGTGAAGGGGAGACATTTTGGCAAGTGCATGAAATGCTAACAGCTG 300
241 GGTGACACTAGTGAAGGGGAGACATTTTGGCAAGTGCATGAAATGCTAACAGCTG 300
301 GCGTTTACACAGAGAGCTTTTCTTAAACAGTATCCCAATGCTGTGACCTGCAAGAACT 360
301 GCGTTTACACAGAGAGCTTTTCTTAAACAGTATCCCAATGCTGTGACCTGCAAGAACT 360
361 GAATTGGAATGTGTAATGCTGTAAGTCTGTCGCGATGATTTCTAACAAATGAGACA 420
361 GAATTGGAATGTGTAATGCTGTAAGTCTGTCGCGATGATTTCTAACAAATGAGACA 420
421 TTACTGTCTCTTAAAGAAAAAATCCACAGCTTCCAGATTAAGTTTCTCAAAATAC 480
421 TTACTGTCTCTTAAAGAAAAAATCCACAGCTTCCAGATTAAGTTTCTCAAAATAC 480
481 ACAAACCTTAAAGATATTTCTTCAAGCAATTTGATTAAGCAATATCCACGAGAAACA 540
481 ACAAACCTTAAAGATATTTCTTCAAGCAATTTGATTAAGCAATATCCACGAGAAACA 540
541 TTTTGTGATTAATGTAATGTAATATTAATATCTCAACCAATGATGATCAACCTTC 600
541 TTTTGTGATTAATGTAATGTAATATTAATATCTCAACCAATGATGATCAACCTTC 600

601 AGACCTGGAATATTCAGAAAGCTTACATGAGCTTAATGCTTAATCTGATGACATCCA 660
601 AGACCTGGAATATTCAGAAAGCTTACATGAGCTTAATGCTTAATCTGATGACATCCA 660
661 ATTAACCAAAATTTTCACAGCGCTTGTGTTTAAACGGGATTTAAATTCCTGTTTCTGTCATG 720
661 ATTAACCAAAATTTTCACAGCGCTTGTGTTTAAACGGGATTTAAATTCCTGTTTCTGTCATG 720
721 GTTAATACTACTAGAGAGCTCTCCCAAGAGATGTGTCCCAATGCTCAACTCAAC 780
721 GTTAATACTACTAGAGAGCTCTCCCAAGAGATGTGTCCCAATGCTCAACTCAAC 780
781 TGGGTGATTTTGAAGCAATAGATAATTAAGTCTCAAAATTCCTGATGCTGCTG 840
781 TGGGTGATTTTGAAGCAATAGATAATTAAGTCTCAAAATTCCTGATGCTGCTG 840
841 GATTGCTCACAGTGTCTGTTTCTGCTTGAAGAAATCAAAATGGTTTGTTCAGAGAAACA 900
841 GATTGCTCACAGTGTCTGTTTCTGCTTGAAGAAATCAAAATGGTTTGTTCAGAGAAACA 900
901 TTTTCTCATTTAAAAAATTTAGAGAACTGATCTGTCTGCAATACGATTAACGAGCTA 960
901 TTTTCTCATTTAAAAAATTTAGAGAACTGATCTGTCTGCAATACGATTAACGAGCTA 960
961 TCACTCACCCTTTTAAAGACTTGAAGCTTCAAAAAGCTGAACCTGATCCAAATCT 1020
961 TCACTCACCCTTTTAAAGACTTGAAGCTTCAAAAAGCTGATCCAAATCT 1020
1021 CTATGATCTTTTCAAGAAACGATTTGAAGTCTTAAACAACTTCAAGTCTTCAAGCTG 1080
1021 CTATGATCTTTTCAAGAAACGATTTGAAGTCTTAAACAACTTCAAGTCTTCAAGCTG 1080
1081 GAAAGATAGAGATTCCAATATTAACACAGAAATTTTCAACCCATGAAGAACTTTTCT 1140
1081 GAAAGATAGAGATTCCAATATTAACACAGAAATTTTCAACCCATGAAGAACTTTTCT 1140
1141 CACATTTATTTAAAAAATTTGATGATCTGCTCAATGCTCCCATGCTCCGAATATGTAAG 1200
1141 CACATTTATTTAAAAAATTTGATGATCTGCTCAATGCTCCCATGCTCCGAATATGTAAG 1200
1201 CCTTGAAGGAGAGGAGCTTTCTTCAATTTAGAGACCTGTTGGCTTAACAATATCTCAGAAATA 1260
1201 CCTTGAAGGAGAGGAGCTTTCTTCAATTTAGAGACCTGTTGGCTTAACAATATCTCAGAAATA 1260
1261 TTTGTCTGGTTATAGCTTTCACTACCTGCTTGAAGAAATCTTTTGTGCAATGGCAATGAGA 1320
1261 TTTGTCTGGTTATAGCTTTCACTACCTGCTTGAAGAAATCTTTTGTGCAATGGCAATGAGA 1320
1321 TCTTTCAATTAAGCTGAAATTAACAACGCTATGTCATCAAAATCTTTGTGTGCT 1380
1321 TCTTTCAATTAAGCTGAAATTAACAACGCTATGTCATCAAAATCTTTGTGTGCT 1380
1381 GATTGCTGATGGAGGTTTAACTGTTCTTTTGGCAATTTTGGATTAATAATCCGAGGG 1440
1381 GATTGCTGATGGAGGTTTAACTGTTCTTTTGGCAATTTTGGATTAATAATCCGAGGG 1440
1441 CAGTATCAGAAATATGCTTCTGCTGATGAGAGAGCTGAGAGCTGATGAGAGG 1500
1441 CAGTATCAGAAATATGCTTCTGCTGATGAGAGAGCTGAGAGCTGATGAGAGG 1500
1501 CTGGCAATGCTGTCAACGAGAGCTCTGTTCTGCTAATGACCTTCTTGAAGAG 1560
1501 CTGGCAATGCTGTCAACGAGAGCTCTGTTCTGCTAATGACCTTCTTGAAGAG 1560
1561 TTCTGTGATATGTTCTTCCCTTCAAGTAACTTGAACCTGGAAGAAACGAGCACTCACTC 1620
1561 TTCTGTGATATGTTCTTCCCTTCAAGTAACTTGAACCTGGAAGAAACGAGCACTCACTC 1620
1621 ATCTCATTTTGCATCTGATGAGCGGAAATTTTAATAGCTTAATTCATTTTGAATTAAG 1680
1621 ATCTCATTTTGCATCTGATGAGCGGAAATTTTAATAGCTTAATTCATTTTGAATTAAG 1680

QY 1681 GATTATTTTGGAACTTTTATGGGAAAAATGAGATGTTTCCACTTATATGACCA 1740
DB 1681 GATTATTTTGGAACTTTTATGGGAAAAATGAGATGTTTCCACTTATATGACCA 1740
QY 1741 ACAGAAAGTATTTGGAAGCAAGGATTTCTCTGGAAATTTTCTTAAAGTGAATCTGCTG 1800
DB 1741 ACAGAAAGTATTTGGAAGCAAGGATTTCTCTGGAAATTTTCTTAAAGTGAATCTGCTG 1800
QY 1801 GCTTTTCTCATCATTTGTTTCTCTATATATATATGTTCTGTTCCATTCAGAAAAACGCC 1860
DB 1801 GCTTTTCTCATCATTTGTTTCTCTATATATATATGTTCTGTTCCATTCAGAAAAACGCC 1860
QY 1861 TTGCAAGACCAAGATGAGAAATGTTTGGAGAGAGATGCTGTTGCAATCGTTTC 1920
DB 1861 TTGCAAGACCAAGATGAGAAATGTTTGGAGAGAGATGCTGTTGCAATCGTTTC 1920
QY 1921 TTTTATATAGTGTCTCTGATGCACTCTGCTGATTTCTGTTATTTGTTAAATCTCTT 1980
DB 1921 TTTTATATAGTGTCTCTGATGCACTCTGCTGATTTCTGTTATTTGTTAAATCTCTT 1980
QY 1981 TCCCTCTCCGGGTGGAAATACAGACACAATGACTTCTGGAATAGTATTTTCTCT 2040
DB 1981 TCCCTCTCCGGGTGGAAATACAGACACAATGACTTCTGGAATAGTATTTTCTCT 2040
QY 2041 CCACTTAAACAGTGTGATGCAATCTCTATATCTCTCAACCACTTTTAAAGAC 2100
DB 2041 CCACTTAAACAGTGTGATGCAATCTCTATATCTCTCAACCACTTTTAAAGAC 2100
QY 2101 AAGTTGAAACAGTGTGCAACCAATCTCTGATATCTCTCAACCACTTTTAAAGAC 2160
DB 2101 AAGTTGAAACAGTGTGCAACCAATCTCTGATATCTCTCAACCACTTTTAAAGAC 2160
QY 2161 AGTTTATCTATCATCTGTTGATGAGAGACTCTCTCTGGAATCTGAGGTTTGG 2220
DB 2161 AGTTTATCTATCATCTGTTGATGAGAGACTCTCTCTGGAATCTGAGGTTTGG 2220
QY 2221 AACCAAAATACACTTGGAGACAGATTAATGAACCAAGTTCTCTAG 2265
DB 2221 AACCAAAATACACTTGGAGACAGATTAATGAACCAAGTTCTCTAG 2265

RESULT 2
AP403384 2838 bp mRNA linear PRI 19-FEB-2002
LOCUS Homo sapiens IGR8 mRNA, complete cds.
DEFINITION AF403384
ACCESSION AF403384
VERSION AF403384.2 GI:18702459
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2838)
AUTHORS Hsu, S.Y., Nakabayashi, K., Nishi, S., Kumagai, J., Kudo, M.,
Sherwood, O.D., and Hsueh, A.-J.
TITLE Activation of orphan receptors by the hormone relaxin
JOURNAL Science 295 (5555), 671-674 (2002)
MEDLINE 21669315
PUBMED 11809971
REFERENCE 2 (bases 1 to 2838)
AUTHORS Hsu, S.Y., Nakabayashi, K., and Bhalla, A.
TITLE Direct Submision
JOURNAL Submitted (26-JUL-2001) GYN/OB, Stanford University, MSOB S385,
Stanford, CA 94305, USA
REFERENCE 3 (bases 1 to 2838)
AUTHORS Hsueh, A.-J.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) Division of Reproductive Biology,
Department of Gynecology and Obstetrics, Stanford University
Medical Center, 300 Pasteur Drive, Room A-344, Stanford, CA
94305-5317, USA
REMARK Sequence update by submitter
COMMENT On Feb 19, 2002 this sequence version replaced gi:18419431.

FEATURES
source
Location/Qualifiers
1..2838
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
107..2371
/codon_start=1
/product="IGR8"
/protein_id="AAU69324.2"
/db_xref="GI:18702460"
/translation="MIVFLVFKHLFSRLITWFEFLHFLIVLNVKDFALVQGSMTIPS
CQGYRPFQCNLTGCLPFAHCDGDKDQNGNDENCGDSGTGATIPGTHGANSVLT
TQSCPLKQYRQCCDCKETECVNGDLKSPMSNNVTLLSLKXKHSLPDKVETLY
TKAKITFLQNCITRHSIRKAPFGICNQLIYLHNCTTTPRPIFDLHULITLID
NPITRISQRLFTGLNSLFLPSVNNVNDLALPKQCAQMPQNLNVDLEGNHLYUTST
PLSCDLTVLFLPRNOIGVPEBKFTSLKNLGLDLSNNTITELSPHLFDLKLQTL
NLSNPLMPLVHKKQFESLQSLDLERIEIPNINTEMPOPMKXLSHLYEKNFYCSY
APHVRIQMPLELDISPELDLANNILRIFWVIAFLTCFQNLNLPVIGRSFLKENTH
AMSKILCCADCLMGVYLPFVGIPIKVRQYOKVALLMMSVOCRLMGFLAMSTEV
SVLLITLYLIERKLVIVFPFSNTRPGRQTSVLICIMAGFLIATVPNNKQFQNF
YKNGVCFPLVYQTEDIGSKYSGLFPLDNLALFLIVFSYITMFCSTOKTALQTT
EVKNCQREVAIVANRFFVIFSDALCWIIPVVKILSLFVBLIPDTMTSMIVIFELV
NSALNPLLYTLNTPFDFDKLKLQHLKHQRKSLFIKIKKSLSTSIWIIEDSSILGLVL
NKITLGSIMKPS"

BASE COUNT 817 a 616 c 536 g 869 t
ORIGIN
Query Match 100.0%; Score 2265; DB 9; Length 2838;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGATTTGTTTTCGTTTAAACATCTCTGACGCTCAGATTGATTAAGTTCTTT 60
DB 107 ATGATTTGTTTTCGTTTAAACATCTCTGACGCTCAGATTGATTAAGTTCTTT 166
QY 61 CTACTTCAATTCATCGTTCTGATCAATGCAAGATTTTGCATGACTCAAGTATGATG 120
DB 167 CTACTTCAATTCATCGTTCTGATCAATGCAAGATTTTGCATGACTCAAGTATGATG 226
QY 121 ATCACTCTTCATGCCAAAAAGATTTTCCCTGAGAACTTACCAAGTCTTACC 180
DB 227 ATCACTCTTCATGCCAAAAAGATTTTCCCTGAGAACTTACCAAGTCTTACC 286
QY 181 CGAGCTTTTACGTGATGGCAAGAGATCTGTGGGAAACGGGCGGACGAAGAACTGT 240
DB 287 CGAGCTTTTACGTGATGGCAAGAGATCTGTGGGAAACGGGCGGACGAAGAACTGT 346
QY 241 GGTGACACTAGTAGGAGGAGCAATATTGGGACAGTGCATGGAATGTCAACAGCGTG 300
DB 347 GGTGACACTAGTAGGAGGAGCAATATTGGGACAGTGCATGGAATGTCAACAGCGTG 406
QY 301 GCCTTAAACAGAGATGCTTTTAAACAGTATTCACAAATGCTGTGACTCAAGAACT 360
DB 407 GCCTTAAACAGAGATGCTTTTAAACAGTATTCACAAATGCTGTGACTCAAGAACT 466
QY 421 TTAAGTGAAGTGAATGCTGATTAAGTCTGTGCGGATGATTTCTAACAATGTGACA 420
DB 467 TTAAGTGAAGTGAATGCTGATTAAGTCTGTGCGGATGATTTCTAACAATGTGACA 526
QY 481 TTAAGTGAAGTGAATGCTGATTAAGTCTGTGCGGATGATTTCTAACAATGTGACA 480
DB 527 TTAAGTGAAGTGAATGCTGATTAAGTCTGTGCGGATGATTTCTAACAATGTGACA 586
QY 541 TTTTGGATTATGATATGCAAAATATTATCTCAACCAAGTGCATCAACCTTC 600
DB 647 TTTTGGATTATGATATGCAAAATATTATCTCAACCAAGTGCATCAACCTTC 706

QY 601 AGACCTGGAATATTCAGAGCTTACATGCTAGCTGCTAATTCAGATGACATCCA 660
 Db 707 AGACCTGGAATATTCAGAGCTTACATGCTAGCTGCTAATTCAGATGACATCCA 766
 QY 661 ATACCAAGATTTTCAAGGCTTTGTTACGGGATTAATTCCTGTTTCTGCTATG 720
 Db 767 ATACCAAGATTTTCAAGGCTTTGTTACGGGATTAATTCCTGTTTCTGCTATG 826
 QY 721 GTTAACTACTAGTGAAGCTCTCCCAAGAGATGCTGAGCAATGCTCAATCCAC 780
 Db 827 GTTAACTACTAGTGAAGCTCTCCCAAGAGATGCTGAGCAATGCTCAATCCAC 886
 QY 781 TGGGAGATTTGGAAGGCAATGAAATTAAGTATCTACAAATTTCAAGCTTTCTGCTG 840
 Db 887 TGGGAGATTTGGAAGGCAATGAAATTAAGTATCTACAAATTTCAAGCTTTCTGCTG 946
 QY 841 GATTGCTCACAGGCTGTTCTGCTGCTAAGAAATCAATTTGTTTGTTCAGAGAAC 900
 Db 947 GATTGCTCACAGGCTGTTCTGCTGCTAAGAAATCAATTTGTTTGTTCAGAGAAC 1006
 QY 901 TTTTCTTATTAATAAATTTAGAGAACTGATCTGCTAGCAATAGATTAAGAGCTA 960
 Db 1007 TTTTCTTATTAATAAATTTAGAGAACTGATCTGCTAGCAATAGATTAAGAGCTA 1066
 QY 961 TCACCTCCTTTTAAAGACTTGAAGCTTCTACAAAGCTGAACCTGCTCAATCCAT 1020
 Db 1067 TCACCTCCTTTTAAAGACTTGAAGCTTCTACAAAGCTGAACCTGCTCAATCCAT 1126
 QY 1021 CTATATGATCTTCAAGAACCTGTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1080
 Db 1127 CTATATGATCTTCAAGAACCTGTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1186
 QY 1081 GAAAGATAGAGATTTCCAAATATATAACAAGAAATGTTTCAACCAAGAGAAATCTTCT 1140
 Db 1187 GAAAGATAGAGATTTCCAAATATATAACAAGAAATGTTTCAACCAAGAGAAATCTTCT 1246
 QY 1141 CACATTTATTTCAAAAATTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 Db 1247 CACATTTATTTCAAAAATTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306
 QY 1201 CCTTGAAGGAGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Db 1307 CCTTGAAGGAGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366
 QY 1261 TTTTCTGAGTATAGCTTTTCAATTAACCTGCTTGAAGATCTTTTGTCTATGAGATGA 1320
 Db 1367 TTTTCTGAGTATAGCTTTTCAATTAACCTGCTTGAAGATCTTTTGTCTATGAGATGA 1426
 QY 1321 TCTTCTATTAAGGCTGAATAATACAACTCAGCTATGCTGCTGCTGCTGCTGCT 1380
 Db 1427 TCTTCTATTAAGGCTGAATAATACAACTCAGCTATGCTGCTGCTGCTGCTGCT 1486
 QY 1381 GATTGCTGAGTATAGCTTTTCAATTAACCTGCTTGAAGATCTTTTGTCTATGAGATGA 1440
 Db 1487 GATTGCTGAGTATAGCTTTTCAATTAACCTGCTTGAAGATCTTTTGTCTATGAGATGA 1546
 QY 1441 CAGTATCAGAAAGTATAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 Db 1547 CAGTATCAGAAAGTATAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606
 QY 1501 CTGGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Db 1607 CTGGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1666
 QY 1561 TTCTGCTGATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 Db 1667 TTCTGCTGATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1726
 QY 1621 ATCCCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 Db 1727 ATCCCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1786
 QY 1681 GATTATTTTGAAGATTTTATGGAAGAAATGAGATATGTTTCCACTTATATGACAA 1740

Db 1787 GATTATTTTGAAGATTTTATGGAAGAAATGAGATATGTTTCCACTTATATGACAA 1846
 QY 1741 ACAGAAATATTTGGAAGGAGATTTCTTGAATTTTCTTGAAGTATGCTGCTGCTG 1800
 Db 1847 ACAGAAATATTTGGAAGGAGATTTCTTGAATTTTCTTGAAGTATGCTGCTGCTG 1906
 QY 1801 GCTTTTCTGATATGTTGTTTCTTGAATTTTCTTGAATTTTCTTGAAGTATGCTGCTG 1860
 Db 1907 GCTTTTCTGATATGTTGTTTCTTGAATTTTCTTGAATTTTCTTGAAGTATGCTGCTG 1966
 QY 1861 TTGAGAACACAGAAATGAGAAATTTGTTGGAAGAGAGGCTGCTGCTGCTGCTGCT 1920
 Db 1967 TTGAGAACACAGAAATGAGAAATTTGTTGGAAGAGAGGCTGCTGCTGCTGCTGCT 2026
 QY 1921 TTTTATATGATGTTCTGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
 Db 2027 TTTTATATGATGTTCTGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2086
 QY 1981 TCCCTCTTCCGGGTGGAATATACAGACATGACTTCTGATAGTATGATTTTCTTCT 2040
 Db 2087 TCCCTCTTCCGGGTGGAATATACAGACATGACTTCTGATAGTATGATTTTCTTCT 2146
 QY 2041 CCAATTAACAGTGTGTTGAATCCAAATCTTATATCTGCTGCTGCTGCTGCTGCTGCT 2100
 Db 2147 CCAATTAACAGTGTGTTGAATCCAAATCTTATATCTGCTGCTGCTGCTGCTGCTGCT 2206
 QY 2101 AAGTTGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 Db 2207 AAGTTGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2266
 QY 2161 AGTTTATCTATCATCTGTTGTTGGAATGAGAGTCTTCTGCTGCTGCTGCTGCTGCT 2220
 Db 2267 AGTTTATCTATCATCTGTTGTTGGAATGAGAGTCTTCTGCTGCTGCTGCTGCTGCT 2326
 QY 2221 AACAAATATACATCTGAGAGCACTATATGAAACAGTTTCTGAG 2265
 Db 2327 AACAAATATACATCTGAGAGCACTATATGAAACAGTTTCTGAG 2371

RESULT 3
 AF453828
 LOCUS
 DEFINITION
 Homo sapiens G protein-coupled receptor affecting testicular
 descent (GRBAT) mRNA, complete cds.
 AF453828
 VERSION
 AF453828.1 GI:18483167
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2436)
 Goriyov, I.P., Kamat, A., Bogatcheva, N.V., Jones, B., Lamb, D.J.,
 Truong, A., Bishop, C.B., McGilreavey, K. and Agoulnik, A.I.
 Mutations of the GRBAT gene cause cryptorchidism
 Hum. Mol. Genet. 11 (19), 2309-2318 (2002)
 12217959
 2 (bases 1 to 2436)
 REFERENCE
 Agoulnik, A.I.
 Direct Submision
 Submitted (30-NOV-2001) Ob/Gyn, Baylor College of Medicine, 6550
 Fannin St., Su. 861, Houston, TX 77030, USA
 FEATURES
 source
 1..2436
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /map="13q12-q13"
 1..2436
 /gene="GREAT"
 71..2335
 CDS

Oy	1801	GCTTTTCTCATCATTTGGTTTTTCTCTATATTA	CTATGTTCTGTTCATTCAAAAAACGGCC	1860
Db	1871	GCTTTTCTCATCATTTGGTTTTTCTCTATATTA	CTATGTTCTGTTCATTCAAAAAACGGCC	1930
Oy	1861	TTGCAAGACCAAGATAGGAATGTGTTTGGAA	GAGAGTGGCGTGTGCAATCGTTTC	1920
Db	1931	TTGCAAGACCAAGATAGGAATGTGTTTGGAA	GAGAGTGGCGTGTGCAATCGTTTC	1990
Oy	1921	TTTTTATAGTGTCTCTGATGCCATCTGCTG	AATTCCTGATATTTGATTAATCCCTT	1980
Db	1991	TTTTTATAGTGTCTCTGATGCCATCTGCTG	AATTCCTGATATTTGATTAATCCCTT	2050
Oy	1981	TCCCTCTTCGGGGTGGAAATACAGACACAT	AGACTTCTCTGTGATAGTATTTTTTCTT	2040
Db	2051	TCCCTCTTCGGGGTGGAAATACAGACACAT	AGACTTCTCTGTGATAGTATTTTTTCTT	2110
Oy	2041	CCAGTTAAACATGCTTTGATATCCATCCTCT	ATATCTTCACAAACCACTTTTAAAGAC	2100
Db	2111	CCAGTTAAACATGCTTTGATATCCATCCTCT	ATATCTTCACAAACCACTTTTAAAGAC	2170
Oy	2101	AAGTTGAAACAGCTGCTGCACAAACATCAG	AGGAATCAATTTTCAAAATTAACAAAAA	2160
Db	2171	AAGTTGAAACAGCTGCTGCACAAACATCAG	AGGAATCAATTTTCAAAATTAACAAAAA	2230
Oy	2161	AGTTTATCTACATCCATCTGTGTGTGATAG	AGACATCTCTTCCCTGAAACTTGGGGTTTG	2220
Db	2231	AGTTTATCTACATCCATCTGTGTGTGATAG	AGACATCTCTTCCCTGAAACTTGGGGTTTG	2290
Oy	2221	AACAAATATACACTTGAGAGACAGTATTA	TGAAACAGATTTCCTAG	2265
Db	2291	AACAAATATACACTTGAGAGACAGTATTA	TGAAACAGATTTCCTAG	2335
RESULT 4				
LOCUS	BD141180	2262 bp	DNA	linear
DEFINITION	Novel G protein-coupled receptor protein and its DNA.			
ACCESSION	BD141180			
VERSION	BD141180.1	GI:23236125		
KEYWORDS	WO 0204640-A/9.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
AUTHORS	1 (bases 1 to 2262)			
TITLE	Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.			
JOURNAL	Novel G protein-coupled receptor protein and its DNA			
COMMENT	Patent: WO 0204640-A 9 17-JAN-2002; TAKEDA CHEMICAL INDUSTRIES LTD, TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA OS Homo sapiens (human) PN WO 0204640-A/9 PD 17-JAN-2002 PE 06-JUL-2001 WO 2001JP005878 PR 06-JUL-2000 JP 00P 211989,18-DEC-2000 JP 00P 383794 PI TAKEO MORIYA,TAKASHI ITO,YASUSHI SHINTANI,NOBUYUKI MIYAJIMA PC C12N15/12,C12N1/21,C07K14/705,C07K16/28,C12P21/02,C1201/68, PC A61K38/00, PC A61K45/00,A61K48/00,A61P1/00,A61P3/00,A61P9/00,A61P25/28, PC A61P29/00, PC A61P31/00,A61P37/00,GO1N33/15,GO1N33/50,GO1N33/53,GO1N33/566// PC (C12N1/21,C12P1.19),(C12P21/02,C12R1.19) CC Novel G protein-coupled receptor protein and its DNA. FH Key Location/Qualifiers 1..2262 FT source FT Location/Qualifiers 1..2262 Location/Qualifiers 1..2262 /organism='Homo sapiens (human)'. /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606' 650 a 467 c 421 g 724 t			
BASE COUNT				

Query Match	99.7%;	Score 2258.8;	DB 6;	Length 2262;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2260;	Conservative	0;	Indels 2;	Gaps 0;
1 ATGATGTTTTCGCGTTTAAACATCTCTCCAGCCGAGATTGATTAACATGTTCTT				60
1 ATGATGTTTTCGCGTTTAAACATCTCTCCAGCCGAGATTGATTAACATGTTCTT				60
61 CTACTTCATTTTCATCGTTCTGATCAATGTCAAAGATTTTGACTGACTCAAGTAGCATG				120
61 CTACTTCATTTTCATCGTTCTGATCAATGTCAAAGATTTTGACTGACTCAAGTAGCATG				120
121 ATCACTCCTTCATGCGCAAAAGAGATTTTTCCTGCGGAAATCTTACCAAGGCTTACC				180
121 ATCACTCCTTCATGCGCAAAAGAGATTTTTCCTGCGGAAATCTTACCAAGGCTTACC				180
181 CGAGCTTTTCACCTGATGCGAAGGATGACTGTGGGAACGGGGCGGACGAAGAGAACTGT				240
181 CGAGCTTTTCACCTGATGCGAAGGATGACTGTGGGAACGGGGCGGACGAAGAGAACTGT				240
241 GGTGCACTAGTGGATGGGCGCAACATATTTGGCACAATGTCATGAAATGCTAACAGCTG				300
241 GGTGCACTAGTGGATGGGCGCAACATATTTGGCACAATGTCATGAAATGCTAACAGCTG				300
301 GCCTTAACACAGGAGTGTCTTCTAAACAGTATCCACAATGCTGACATGCGCAAGAACT				360
301 GCCTTAACACAGGAGTGTCTTCTAAACAGTATCCACAATGCTGACATGCGCAAGAACT				360
361 GAAATGGAATGTGTAATGATGACTTAAAGTCTGCGGATGATTTCTAACATGTGACA				420
361 GAAATGGAATGTGTAATGATGACTTAAAGTCTGCGGATGATTTCTAACATGTGACA				420
421 TTACTGTCTCTTAAGAAAAAATAATCCACAGTCTTCCAGATTAAGTTTCTATCAATAC				480
421 TTACTGTCTCTTAAGAAAAAATAATCCACAGTCTTCCAGATTAAGTTTCTATCAATAC				480
481 ACAAAACCTTAAAAAAGATATTTCTTCAGACATAATTCATTAGACATAATCCAGAAAGCA				540
481 ACAAAACCTTAAAAAAGATATTTCTTCAGACATAATTCATTAGACATAATCCAGAAAGCA				540
541 TTTTGTGATTAATGTAATCTGCAATATTAATATCTCAACCAACATCTGATCAACACCTC				600
541 TTTTGTGATTAATGTAATCTGCAATATTAATATCTCAACCAACATCTGATCAACACCTC				600
601 AGACCTGGAATTTCAAAAGACTTAATCATGAGCTAACTGGCTTAATCTTAGATGACAATCA				660
601 AGACCTGGAATTTCAAAAGACTTAATCATGAGCTAACTGGCTTAATCTTAGATGACAATCA				660
661 ATTAACAGAAATTTCAACAGGCTTGTTAACGGGATTAATTCCTGTTTTCTGTCTATG				720
661 ATTAACAGAAATTTCAACAGGCTTGTTAACGGGATTAATTCCTGTTTTCTGTCTATG				720
721 GTTAATACTACTTTAGAACTCTTCCCAAGAGAGTGGCCCAATGCTCAACTCAAC				780
721 GTTAATACTACTTTAGAACTCTTCCCAAGAGAGTGGCCCAATGCTCAACTCAAC				780
781 TGGGTGGAATTTGGAAGGCAATAGAAATTAAGTATCTCACAAAATCTCACTGTTCTGTG				840
781 TGGGTGGAATTTGGAAGGCAATAGAAATTAAGTATCTCACAAAATCTCACTGTTCTGTG				840
841 GATTGGCTCACAGTGTGTTTCTGCTAGAAAATCAAAATGGTTTTTGTTCAGAGAAAGCA				900
841 GATTGGCTCACAGTGTGTTTCTGCTAGAAAATCAAAATGGTTTTTGTTCAGAGAAAGCA				900
901 TTTTCTCATTAATAAAATTTAGAGAACTGGAATCTGTACGAAATACGATTAACGGAAC				960
901 TTTTCTCATTAATAAAATTTAGAGAACTGGAATCTGTACGAAATACGATTAACGGAAC				960
961 TCACTCACTTTTAAAGACTTGAAGCTTCAACAAAGCTGAACCTGTATCAATCTCT				1020
961 TCACTCACTTTTAAAGACTTGAAGCTTCAACAAAGCTGAACCTGTATCAATCTCT				1020

QY 1021 CTATGATCTTCAAGAACCAAGTTGAAGCTTAAACAATCTAGCTCTAGACCTG 1080
 Db 1021 CTATGATCTTCAAGAACCAAGTTGAAGCTTAAACAATCTAGCTCTAGACCTG 1080
 QY 1081 GAAAGATAGATTCACAAATATAACACAGAAATGTTTCAACCCATGAAAGATCTTCT 1140
 Db 1081 GAAAGATAGATTCACAAATATAACACAGAAATGTTTCAACCCATGAAAGATCTTCT 1140
 QY 1141 CACATTTATTTCAAAAACCTTGATAGCTCTATAGTCCCAATGTCGGAATATGATG 1200
 Db 1141 CACATTTATTTCAAAAACCTTGATAGCTCTATAGTCCCAATGTCGGAATATGATG 1200
 QY 1201 CCCTTGACGACGACGATTTCTTCAATTGAGGACCTCTTGCTAAACAATATCTCGAATA 1260
 Db 1201 CCCTTGACGACGACGATTTCTTCAATTGAGGACCTCTTGCTAAACAATATCTCGAATA 1260
 QY 1261 TTTGCTGAGGTTATAGCTTCAATACCTGCTTGGAAATCTTTTGTGATGGCATGAGA 1320
 Db 1261 TTTGCTGAGGTTATAGCTTCAATACCTGCTTGGAAATCTTTTGTGATGGCATGAGA 1320
 QY 1321 TCTTTCATTAAGCTGAATAATACCTGATAGTGCATCAAAAATCTTGTGTGCT 1380
 Db 1321 TCTTTCATTAAGCTGAATAATACCTGATAGTGCATCAAAAATCTTGTGTGCT 1380
 QY 1381 GATTGCGCTGATGGGTGTTTACTTGTCTTGTGTCATTTTGCATATATAAAATACCGAGG 1440
 Db 1381 GATTGCGCTGATGGGTGTTTACTTGTCTTGTGTCATTTTGCATATATAAAATACCGAGG 1440
 QY 1441 CAGATACGAAATATGCTTGTCTGATGAGACCGGACAGTGGCGCTCATGGGGTTC 1500
 Db 1441 CAGATACGAAATATGCTTGTCTGATGAGACCGGACAGTGGCGCTCATGGGGTTC 1500
 QY 1501 CTGGCAGTGTCTGACCAAGGATCTGTGCTGCTACTGACCTACTGATCTTGGAGAG 1560
 Db 1501 CTGGCAGTGTCTGACCAAGGATCTGTGCTGCTACTGACCTACTGATCTTGGAGAG 1560
 QY 1561 TTTCTGCTGATTTGCTTCCCTTCAATTAACATTCGACCTGGAACCGGACAGCTCAGTC 1620
 Db 1561 TTTCTGCTGATTTGCTTCCCTTCAATTAACATTCGACCTGGAACCGGACAGCTCAGTC 1620
 QY 1621 ATCTCATTTGCAATCTGATAGTGGGAGATTTTAAATAGTCTATATTCATTTTGAATAG 1680
 Db 1621 ATCTCATTTGCAATCTGATAGTGGGAGATTTTAAATAGTCTATATTCATTTTGAATAG 1680
 QY 1681 GATTATTTTGGAAACCTTTATGAGGAAATGAGATGTTCCCACTTTATATGACCA 1740
 Db 1681 GATTATTTTGGAAACCTTTATGAGGAAATGAGATGTTCCCACTTTATATGACCA 1740
 QY 1741 ACAGAAATATTTGGAAGCAAGGATTTCTTGAATTTTCTAGTGTGAACTTGTG 1800
 Db 1741 ACAGAAATATTTGGAAGCAAGGATTTCTTGAATTTTCTAGTGTGAACTTGTG 1800
 QY 1801 GCTTTTTCATCATTTGTTTCCATATATATATGTTCTGTTCATTTCAAAAACCGCC 1860
 Db 1801 GCTTTTTCATCATTTGTTTCCATATATATATGTTCTGTTCATTTCAAAAACCGCC 1860
 QY 1861 TTTGAGACCAAGAAATGTTTGAAGAGAGTGGCTGTGCAATGCTTTC 1920
 Db 1861 TTTGAGACCAAGAAATGTTTGAAGAGAGTGGCTGTGCAATGCTTTC 1920
 QY 1921 TTTTATATAGTGTCTGATGCAATCTGCTGATTTCTGTATTTGATTAATCTT 1980
 Db 1921 TTTTATATAGTGTCTGATGCAATCTGCTGATTTCTGTATTTGATTAATCTT 1980
 QY 1981 TCCCTCTCCGGGTGGAATACCAACAATGATCTTCTGATGATGATTTTCTT 2040
 Db 1981 TCCCTCTCCGGGTGGAATACCAACAATGATCTTCTGATGATGATTTTCTT 2040
 QY 2041 CCAAGTTAAAGTCTTTGATTCATCTCTATATCTTCAACAACCTTTTAAAGAC 2100
 Db 2041 CCAAGTTAAAGTCTTTGATTCATCTCTATATCTTCAACAACCTTTTAAAGAC 2100

QY 2101 AAGTTGAACAGCTGCTGCAAAACATCAGAGGAAATCAATTTTCAAAATTAATAAAAAA 2160
 Db 2101 AAGTTGAACAGCTGCTGCAAAACATCAGAGGAAATCAATTTTCAAAATTAATAAAAAA 2160
 QY 2161 AGTTATCTACATTCATTTGTGTGATAGAGACTCTCCCTTCCCTGAAACTTTGGGTTTG 2220
 Db 2161 AGTTATCTACATTCATTTGTGTGATAGAGACTCTCCCTTCCCTGAAACTTTGGGTTTG 2220
 QY 2221 AACAAATTAACACTTGGAGACAGTATATGAAACAGTTTCC 2262
 Db 2221 AACAAATTAACACTTGGAGACAGTATATGAAACAGTTTCC 2262
 RESULT 5
 AX451562
 LOCUS
 DEFINITION
 Sequence 1 from Patent WO0226824.
 AX451562
 VERSION
 AX451562.1 GI:21698547
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 Feder, J.N., Mintier, G., Ramanathan, C.S., and Hawken, D.R.
 A novel human g-protein coupled receptor, hgrbm5, expressed
 highly in brain and ovarian tissues
 Patent: WO 0226824-A 1 04-APR-2002;
 JOURNAL
 Bristol-Myers Squibb Company (US)
 Location/Qualifiers
 FEATURES
 source
 1..2214
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT
 638 a 457 c 416 g 703 t
 ORIGIN
 Query Match 97.7%; Score 2214; DB 6; Length 2214;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 52 ATGTTCTTTTACTTCAATTTTCATGCTTGTGATCAATGTCAAAGATTTTGCATGACTCAA 111
 Db 1 ATGTTCTTTTACTTCAATTTTCATGCTTGTGATCAATGTCAAAGATTTTGCATGACTCAA 60
 QY 112 GGTAGCATGATCACTCTTCAATGCGCAAAAGATTTTCCGTGTGGAAATCTTAACAAG 171
 Db 61 GGTAGCATGATCACTCTTCAATGCGCAAAAGATTTTCCGTGTGGAAATCTTAACAAG 120
 QY 172 TGCTTACCCGAGCTTTTCACTGATGAGCAAGGATGCTGGGAAACGGGCGGAGCAA 231
 Db 121 TGCTTACCCGAGCTTTTCACTGATGAGCAAGGATGCTGGGAAACGGGCGGAGCAA 180
 QY 232 GAGAACTGTGTGACACTAGTGGGCGAACATATTTGGCACAGTGCATGAAATGCT 291
 Db 181 GAGAACTGTGTGACACTAGTGGGCGAACATATTTGGCACAGTGCATGAAATGCT 240
 QY 292 AAGAGGTGGCTTAAACAGAGAGTGTCTTCTAATAAGTATCCAAATGCTGAGCTGC 351
 Db 241 AAGAGGTGGCTTAAACAGAGAGTGTCTTCTAATAAGTATCCAAATGCTGAGCTGC 300
 QY 352 AAGAAATGAAATGGAATGTAATGTGACTTAAAGTCTGCGAGTATTTCTAAC 411
 Db 301 AAGAAATGAAATGGAATGTAATGTGACTTAAAGTCTGCGAGTATTTCTAAC 360
 QY 412 AATGTGACTTACCTGCTCTTAAAGAAACAAATCCACAGTCTTCCAGATTAAGTTTG 471
 Db 361 AATGTGACTTACCTGCTCTTAAAGAAACAAATCCACAGTCTTCCAGATTAAGTTTG 420
 QY 472 ATCAATATACCAAAAGTAAAGATATTTCTTCAAGATATGATTAAGACATATCC 531
 Db 421 ATCAATATACCAAAAGTAAAGATATTTCTTCAAGATATGATTAAGACATATCC 480

QY 532 AGAAGAGATTTTGGATTATGTATCTGCAAAATATATATCTCAACCAACATCTGCATC 591
 DB 481 AGGAAAGCATTTTGGATTATGTATCTGCAAAATATATATCTCAACCAACATCTGCATC 540
 QY 592 ACAACCCCTGAGACCTGGAAATATTCAGAAAGCTTACATCAGCTAACTTGGCTAATTCAGAT 651
 DB 541 ACAACCCCTGAGACCTGGAAATATTCAGAAAGCTTACATCAGCTAACTTGGCTAATTCAGAT 600
 QY 652 GACAACTCAATTAACAGAAATTCACAGCGCTGTGTTACGGGATTAATTCCTGTGTTTC 711
 DB 601 GACAACTCAATTAACAGAAATTCACAGCGCTGTGTTACGGGATTAATTCCTGTGTTTC 660
 QY 712 CTGTCTATGGTATTAATCTACTTAAAGCTCTTCCAGAGCATGTGTGCCAAATGCCT 771
 DB 661 CTGTCTATGGTATTAATCTACTTAAAGCTCTTCCAGAGCATGTGTGCCAAATGCCT 720
 QY 772 CAACCTCACTGGGTGGATTGGAAAGCAATAGAAATTAAGTATCTCAAAATTCAGCTT 831
 DB 721 CAACCTCACTGGGTGGATTGGAAAGCAATAGAAATTAAGTATCTCAAAATTCAGCTT 780
 QY 832 CTGTCTGCGATTCGCTCAACAGTCTGTGTTCTGCTAGAAATCAAAATGGTGTGTTCCA 891
 DB 781 CTGTCTGCGATTCGCTCAACAGTCTGTGTTCTGCTAGAAATCAAAATGGTGTGTTCCA 840
 QY 892 GAGAGACATTTCTTCATTAAAAAATTTAGAGAACTGGATCTGTCTAGCAATACGAT 951
 DB 841 GAGAGACATTTCTTCATTAAAAAATTTAGAGAACTGGATCTGTCTAGCAATACGAT 900
 QY 952 AGGAGCTATCACTCACTCTTTTAAAGCTTAAAGCTTACAAAGCTGAACTGTCA 1011
 DB 901 AGGAGCTATCACTCACTCTTTTAAAGCTTAAAGCTTACAAAGCTGAACTGTCA 960
 QY 1012 TCCAACTCTTATGTATCTTCAACAAGACAGTTGAAAGCTTAAACAATCTGACT 1071
 DB 961 TCCAACTCTTATGTATCTTCAACAAGACAGTTGAAAGCTTAAACAATCTGACT 1020
 QY 1072 CTAGACCTGGAAGAGATAGATTCCAAAATTAACAACAGAAATGTTCAACCACTGAA 1131
 DB 1021 CTAGACCTGGAAGAGATAGATTCCAAAATTAACAACAGAAATGTTCAACCACTGAA 1080
 QY 1132 AATCTTCTCACAATTTATTTCAAAAACTTGCATGCTCTATGCTCTATGCTCCCATGTCGA 1191
 DB 1081 AATCTTCTCACAATTTATTTCAAAAACTTGCATGCTCTATGCTCTATGCTCCCATGTCGA 1140
 QY 1192 AATGATATGCCCTTGAAGAGAGGCAATTTCTCATTTGAGGACCTCTTGGCTAAACATATC 1251
 DB 1141 AATGATATGCCCTTGAAGAGAGGCAATTTCTCATTTGAGGACCTCTTGGCTAAACATATC 1200
 QY 1252 CTGAGAAATTTGTCTGGGTTATAGCTTCAATTAACCTGCTTTGGAAATCTTTTGTGACT 1311
 DB 1201 CTGAGAAATTTGTCTGGGTTATAGCTTCAATTAACCTGCTTTGGAAATCTTTTGTGACT 1260
 QY 1312 GGCATGAGATCTTTCATTAAAGCTGAAATTAACAACCTGATGTCATCAAAATCTCT 1371
 DB 1261 GGCATGAGATCTTTCATTAAAGCTGAAATTAACAACCTGATGTCATCAAAATCTCT 1320
 QY 1372 TGTGTGCTGATGCTCTGATGGGTGTTACTGTGTTCTTTGGCAATTTTCGATATTA 1431
 DB 1321 TGTGTGCTGATGCTCTGATGGGTGTTACTGTGTTCTTTGGCAATTTTCGATATTA 1380
 QY 1432 TACCGAGGGAGATTAAGAAATGAGCTGCTGAGATGAGAGAGCGGAGTGCAGCTC 1491
 DB 1381 TACCGAGGGAGATTAAGAAATGAGCTGCTGAGATGAGAGAGCGGAGTGCAGCTC 1440
 QY 1492 ATGGGATTCCTGGCATGCTGTCCACCGAAGCTCTGTGTTCTGCTACTGACCTACTGACT 1551
 DB 1441 ATGGGATTCCTGGCATGCTGTCCACCGAAGCTCTGTGTTCTGCTACTGACCTACTGACT 1500
 QY 1552 TTGAGAAAGTTCTGGGTCATTTGTCTCCCTTCAGTAACTTGAACCTGGAAAAACGGCAG 1611
 DB 1501 TTGAGAAAGTTCTGGGTCATTTGTCTCCCTTCAGTAACTTGAACCTGGAAAAACGGCAG 1560

QY 1612 ACCTGATCATCTCATTTTGCATCTGATTCGCGGGATTTTAAATGCTGTAATTCGATTT 1671
 DB 1561 ACCTGATCATCTCATTTTGCATCTGATTCGCGGGATTTTAAATGCTGTAATTCGATTT 1620
 QY 1672 TGGAAATTAAGGATTTATTTTGGAAAATTTATGGAAGAAAATGAGTATGTTTCCACTTTAT 1731
 DB 1621 TGGAAATTAAGGATTTATTTTGGAAAATTTATGGAAGAAAATGAGTATGTTTCCACTTTAT 1680
 QY 1732 TATGACCAAAACAGAAATATTGGAAGCAAAAGGATATCTCTTGAATTTTCTAGGTGTG 1791
 DB 1681 TATGACCAAAACAGAAATATTGGAAGCAAAAGGATATCTCTTGAATTTTCTAGGTGTG 1740
 QY 1792 AACTGCTGGCTTTCTCATCATTTGATTTTCTTATTAATGATGCTGCTCAATGCA 1851
 DB 1741 AACTGCTGGCTTTCTCATCATTTGATTTTCTTATTAATGATGCTGCTCAATGCA 1800
 QY 1852 AAAACCGCTTGCAGACCAACAGAAATGGAATTTGTTTGAAGAGAGTGTGCTGTGCA 1911
 DB 1801 AAAACCGCTTGCAGACCAACAGAAATGGAATTTGTTTGAAGAGAGTGTGCTGTGCA 1860
 QY 1912 AATGCTTTCTTTTATTAATGATGTTCTGTGATGCTGATCTGCTGATTTCTGATTT 1971
 DB 1861 AATGCTTTCTTTTATTAATGATGTTCTGTGATGCTGATCTGCTGATTTCTGATTT 1920
 QY 1972 AAAATCCCTTCCCTCTCCGGGTGGAATTAACAGACAAATGACTTCCGATGATGATT 2031
 DB 1921 AAAATCCCTTCCCTCTCCGGGTGGAATTAACAGACAAATGACTTCCGATGATGATT 1980
 QY 2032 TTTTTCCTTCCAGTTTAAACAGTCTTGAATCCATCTCTATATCTCAACCAACTTT 2091
 DB 1981 TTTTTCCTTCCAGTTTAAACAGTCTTGAATCCATCTCTATATCTCAACCAACTTT 2040
 QY 2092 TTTTAAAGCAAGTTGAAACAGCTGCTGCAAAACATCAGAGGAATCAATTTTCAAAAT 2151
 DB 2041 TTTTAAAGCAAGTTGAAACAGCTGCTGCAAAACATCAGAGGAATCAATTTTCAAAAT 2100
 QY 2152 AAAAAAAGTTTATCTATCATCTCAATGATGATGATGAGAGCTCCCTCCCTGAAACTT 2211
 DB 2101 AAAAAAAGTTTATCTATCATCTCAATGATGATGATGAGAGCTCCCTCCCTGAAACTT 2160
 QY 2212 GGGGTTTTGAACAAATTAACCTTGGAGACAGTATTAATGAAACAGTTTCTTAA 2265
 DB 2161 GGGGTTTTGAACAAATTAACCTTGGAGACAGTATTAATGAAACAGTTTCTTAA 2214

RESULT 6
 AX385032
 LOCUS 2193 bp DNA linear PAT 19-MAR-2002
 DEFINITION Sequence 6 from Patent WO0214489.
 ACCESSION AX385032
 VERSION AX385032.1 GI:19578156
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
 Leucine-rich repeat-containing G-protein coupled receptor-8
 molecules and uses thereof
 Patent: WO 0214489-A 6 21-FEB-2002;
 JOURNAL
 Amgen, Inc. (US)
 FEATURES
 source
 1..2193 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 1..2193
 CDS
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28588.1"
 /db_xref="GI:19578157"
 /translation="MIVFLVFLHFLSLRLITNPLFLHFTVLINVDPALTGSMITPS

COKEJFPCGJLKLCPRAHOCBGKDDOCAPBESNENGDPSGAMTITGTHYGNANVY
TOEJCKLOPODCCDCTEBCNGBQKSYPAISNNVTLISLKSJGHTSLHPDXFYFIC
TKKJCLPJDOHNCITRHSRAFLRQJCNQJLYLHANCITTLRPGIFKJOLHJWMLDDB
NFTISISQJLPLGJLNSLFLPSLWANNLEALPKQWCOMPLANNVJDEGNISXYLWNS
PLSCSLVJLPLGJLNSLTLSPILFKJOLKJLQJLNSLWANNLEALHROKPSJLQJLST
DLERIEPIINRTPMOPKMSLSTIKYFENKFTYSGASAPVHRCMELTGOISFEDJLANN
LIRIVWVAJATFCEPMLFVJIGRSTIKKENTHANSIKILCDAGLSEVJLFPVGFJ
PIKTSOGVJLWALLMANSOCVRLMGLMSTSVSLVJLJLTXERIVJVPSPNSR
PEKRTOSVJLJLIMAGJLWJLPLMKNYFCNGVSLVJLNGVJLJLTXERIVJVPSPNSR
JGJPLGJLVNLAJLIVSFTITMCSQJLKLQTTBEBRANGREYVANAARFJVFSDA
ICMIFVJFVJLSLSEVJLJLPTWNSLJLFLVNSNALNKLJLTVTLNPFKDKJLQJL
HJHOKSVPKIKKSLSTSVJLWJEDSSJLFLVNSNALNKLJLTVTLNPFKDKJLQJL

sig_peptide	1.	.108
BASE COUNT	626 a	456 c 411 g 700 t
ORIGIN		

Query Match	93.2%	Score 211;	DB 6;	Length 2193;
Best Local Similarity	96.8%	Pred. No. 0;		
Matches 2193; Conservative	0;	Mismatches	0;	Indels 72; Gaps 1;

QY	1	ATGATGTTTTCGTGTTTTTAAACATCTCTTCAAGCTTCAGATTGATTAACAATGTTCTTT	60
Db	1	ATGATGTTTTCGTGTTTTTAAACATCTCTTCAAGCTTCAGATTGATTAACAATGTTCTTT	60
QY	61	CTACTTCATTTGATGCTTGTGATGCATGTCAAAGATTTTGTGACCTGACCTCAAGTGCATG	120
Db	61	CTACTTCATTTGATGCTTGTGATGCATGTCAAAGATTTTGTGACCTGACCTCAAGTGCATG	120
QY	121	ATCACTCCTTCATGCCAAAAAGATATTTTCCCTGTGGAAATCTTAACAAATGCTTAACC	180
Db	121	ATCACTCCTTCATGCCAAAAAGATATTTTCCCTGTGGAAATCTTAACAAATGCTTAACC	180
QY	181	CGAGCTTTTCACTGTGATGGCAAGATGACTGTGGAAACGGGGCGGACGAAAGAACTGT	240
Db	181	CGAGCTTTTCACTGTGATGGCAAGATGACTGTGGAAACGGGGCGGACGAAAGAACTGT	240
QY	241	GGTGCACACTAGTGAATGGCGCAACATATTGGCAGACGTGCAATGTAAGTCTTAACAGCGTG	300
Db	241	GGTGCACACTAGTGAATGGCGCAACATATTGGCAGACGTGCAATGTAAGTCTTAACAGCGTG	300
QY	301	GCCTTAACACAGGAGTGCCTTTCTTAAACAGTATCCAAATGCTGTGACTGCAAGAAACT	360
Db	301	GCCTTAACACAGGAGTGCCTTTCTTAAACAGTATCCAAATGCTGTGACTGCAAGAAACT	360
QY	361	GAATTGGAATGTGTAAATGGTGACTTAAAGTGTGGCGGATGATTTCTTAACAAATGTGACA	420
Db	361	GAATTGGAATGTGTAAATGGTGACTTAAAGTGTGGCGGATGATTTCTTAACAAATGTGACA	420
QY	421	TTACTGTCTCTTAAGAAAAACAAAATCCACAGCTCTTCCAGATTAAGTTTTCATCAAAATAC	480
Db	421	TTACTGTCTCTTAAGAAAAACAAAATCCACAGCTCTTCCAGATTAAGTTTTCATCAAAATAC	480
QY	481	ACAAAACCTTAAAAAGATATTTCTTCAGCAATATTCATTAGACACATATCCAGGAAACGA	540
Db	481	ACAAAACCTTAAAAAGATATTTCTTCAGCAATATTCATTAGACACATATCCAGGAAACGA	540
QY	541	TTTTTGGATTTATGTATCTGCAAAATATATATCTCAACCAACATGCAATCAACCCCTC	600
Db	541	TTTTTGGATTTATGTATCTGCAAAATATATATCTCAACCAACATGCAATCAACCCCTC	600
QY	601	AGACCTGGAATATTCAGAACTTACATCAGCTTAATCTGGCTAATTTCTAGATGACATCCA	660
Db	601	AGACCTGGAATATTCAGAACTTACATCAGCTTAATCTGGCTAATTTCTAGATGACATCCA	660
QY	661	ATTAACCAAAATTTCAAGGCTTGTTTACGGGATTTAAATCTTGTGTTTTCTGTCTATAG	720
Db	661	ATTAACCAAAATTTCAAGGCTTGTTTACGGGATTTAAATCTTGTGTTTTCTGTCTATAG	720
QY	721	GTTATTAATCTACTTGAAGACTCTTCCCAAGACAGATGTGTGCCMAATGCTCTTAATCTCAC	780
Db	721	GTTATTAATCTACTTGAAGACTCTTCCCAAGACAGATGTGTGCCMAATGCTCTTAATCTCAC	780
QY	781	TGGGTGATTTGGAAAGCAATAGAAATTAAGTATCTCAAAATTTCAAGTTTTCTGTCTGTGC	840

[illegible]

Db 1789 TTGCAGACCAAGAGTAAGATTTGTTTGAAGAGAGTGCGTGTGCAATCGTTTC 1848
Qy 1921 TTTTATATAGTGTCTCTGATGCGCATCTGCTGATTCCTGATTTGTAATAATCCTT 1980
Db 1849 TTTTATATAGTGTCTCTGATGCGCATCTGCTGATTCCTGATTTGTAATAATCCTT 1908
Qy 1981 TCCCTCTTCCGGGTGGAATAATACAGACACAATAGACTTCCGTGATGATTTTTCCTT 2040
Db 1909 TCCCTCTTCCGGGTGGAATAATACAGACACAATAGACTTCCGTGATGATTTTTCCTT 1968
Qy 2041 CCAGTTAACAGTGTCTTGAATCCATCTCTATATCTGCAACAACATTTTAAAGAC 2100
Db 1969 CCAGTTAACAGTGTCTTGAATCCATCTCTATATCTGCAACAACATTTTAAAGAC 2028
Qy 2101 AAGTTAAACAGTGTCTGCAACAACATCTAGAGAAATTCATTTTCAAAATTTAAAAA 2160
Db 2029 AAGTTAAACAGTGTCTGCAACAACAATCTAGAGAAATTCATTTTCAAAATTTAAAAA 2088
Qy 2161 AGTTTATCTACATCCATTTGTGTGATAGAGACTCCCTCTCCCTGAAACTTGGGGTTTG 2220
Db 2089 AGTTTATCTACATCCATTTGTGTGATAGAGACTCCCTCTCCCTGAAACTTGGGGTTTG 2148
Qy 2221 AACCAATAACACTTGGAGACAGATATATGAAAACAGTTTCTAG 2265
Db 2149 AACCAATAACACTTGGAGACAGATATATGAAAACAGTTTCTAG 2193
RESULT 7
BD141181 2190 bp DNA linear PAR 18-SEP-2002
LOCUS Novel G protein-coupled receptor protein and its DNA.
DEFINITION BD141181
ACCESSION BD141181 GI:23236126
VERSION WO 0204640-A/10.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2190)
AUTHORS Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
TITLE Novel G protein-coupled receptor protein and its DNA
JOURNAL Patent: WO 0204640-A 10 17-JAN-2002;
TAKEDA CHEMICAL INDUSTRIES LTD, TAKEDA MORIYA, TAKASHI ITO, YASUSHI
SHINTANI, NOBUYUKI MIYAJIMA
COMMENT OS Homo sapiens (human)
PN WO 0204640-A/10
FD 17-JAN-2002
PF 06-JUL-2001 WO 2001JP005878
PR 07-JUL-2000 JP 00P 211989,18-DEC-2000 JP 00P 383794 PI
TAKEDA MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA PC
C12N15/12, C12N1/21, C07K14/705, C07K16/28, C12P21/02, C12Q1/68, PC
A61K38/00,
PC A61K45/00, A61K48/00, A61P1/00, A61P3/00, A61P9/00, A61P25/28, PC
A61P39/00,
PC A61P35/00, A61P37/00, G01N33/15, G01N33/50, G01N33/53, G01N33/566//
PC (C12N1/21, C12R1.19), (C12P21/02, C12R1.19)
CC Novel G protein-coupled receptor protein and its DNA. FH Key
FT source 1. 2190
FT location/Qualifiers
FEATURES
source 1. 2190
location/Qualifiers
1. 2190
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 626 a 457 c 409 g 698 t
ORIGIN
Query Match 92.9%; Score 2104.8; DB 6; Length 2190;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 2188; Conservative 0; Mismatches 2; Indels 72; Gaps 1;
Qy 1 ATGATGTTTTCGTGTTTAAACATCTTTCAGCTCAGATTGATTACATGTTCTTT 60

Db 1 ATGATGTTTTCGTGTTTAAACATCTTTCAGCTCAGATTGATTACATGTTCTTT 60
Qy 61 CTACTTCATTTTCATGTTCTGATGCAATGCAAGATTTTTCATCTGATGATGATG 120
Db 61 CTACTTCATTTTCATGTTCTGATGCAATGCAAGATTTTTCATCTGATGATGATG 120
Qy 121 ATCACTCCCTCAATGCAAGATTTTTCATGATGATGATGATGATGATGATGATG 180
Db 121 ATCACTCCCTCAATGCAAGATTTTTCATGATGATGATGATGATGATGATGATG 180
Qy 181 CGAGCTTTTCACTGATGATGCAAGATGATGATGATGATGATGATGATGATGATG 240
Db 181 CGAGCTTTTCACTGATGATGCAAGATGATGATGATGATGATGATGATGATGATG 240
Qy 241 GGTGACACTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 GGTGACACTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 GCTTAAACAGAGAGGCTTTCTAAACAGATGATGATGATGATGATGATGATGATG 360
Db 301 GCTTAAACAGAGAGGCTTTCTAAACAGATGATGATGATGATGATGATGATGATG 360
Qy 361 GAATGGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Db 361 GAATGGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Qy 421 TTAATGCTCTTAAAGAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 480
Db 421 TTAATGCTCTTAAAGAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 480
Qy 481 TTAATGCTCTTAAAGAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 480
Db 481 TTAATGCTCTTAAAGAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 480
Qy 481 TTAATGCTCTTAAAGAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 480
Db 481 TTAATGCTCTTAAAGAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 480
Qy 541 TTTTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 TTTTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 601 AAGCTCGAATTTTCAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 660
Db 601 AAGCTCGAATTTTCAAAACCAAAATCCAGTCTCCAGATTAAGTTTCTCAATATG 660
Qy 661 ATACAGAAATTTCAAGAGGCTTTGATGATGATGATGATGATGATGATGATGATG 720
Db 661 ATACAGAAATTTCAAGAGGCTTTGATGATGATGATGATGATGATGATGATGATG 720
Qy 721 GTTAAATTAATTAAGAGCTTTCCCAAGAGATGATGATGATGATGATGATGATGATG 780
Db 721 GTTAAATTAATTAAGAGCTTTCCCAAGAGATGATGATGATGATGATGATGATGATG 780
Qy 781 TGGGTGATTTTGAAGCAATGAAATTAAGTATCTCAAAATTTCTGATGATGATGATG 840
Db 781 TGGGTGATTTTGAAGCAATGAAATTAAGTATCTCAAAATTTCTGATGATGATGATG 840
Qy 841 GATTCCTCAAGTGTGTTTCTGCTAGAAATCAATGTTTGTTCAGAGAGAGACA 900
Db 841 GATTCCTCAAGTGTGTTTCTGCTAGAAATCAATGTTTGTTCAGAGAGAGACA 900
Qy 901 TTTTCTTCAATTAATAATTTAGAGAACTGATGATGATGATGATGATGATGATGATG 960
Db 901 TTTTCTTCAATTAATAATTTAGAGAACTGATGATGATGATGATGATGATGATGATG 960
Qy 961 TCACCTCACTTTTAAAGCTTGAAGCTTTCAAAAGCTTTCAAAAGCTTTCAAAAGCTT 1020
Db 961 TCACCTCACTTTTAAAGCTTGAAGCTTTCAAAAGCTTTCAAAAGCTTTCAAAAGCTT 1020
Qy 1021 CTATGATCTTCAAGAGAAACAGTTGAAAGTTTAAACATGATGATGATGATGATGATG 1080
Db 1021 CTATGATCTTCAAGAGAAACAGTTGAAAGTTTAAACATGATGATGATGATGATGATG 1080
Qy 1081 GAAAGATTAAGATTTCAAAATTAATAACAGAGATTTTCAACCATGAAGATCTTTCT 1140
Db 1081 GAAAGATTAAGATTTCAAAATTAATAACAGAGATTTTCAACCATGAAGATCTTTCT 1140

Db 1009 GAAAGATAGAGATTCAGAAATATAACACAGAAATGTTCAACCCATGAGAGATCTTCT 1068
Qy 1141 CACATTTATTTCAAAAACCTTGCATAGTCTCTTATGCTCCCATGTCGAAATATGTATG 1200
Db 1069 CACATTTATTTCAAAAACCTTGCATAGTCTCTTATGCTCCCATGTCGAAATATGTATG 1128
Qy 1201 CCCTTGACGAGCGGATTTCTTCAATTGAGACCTCTTGCTAACAAATATCCTCGAATA 1260
Db 1129 CCCTTGACGAGCGGATTTCTTCAATTGAGACCTCTTGCTAACAAATATCCTCGAATA 1188
Qy 1261 TTTGCTGAGGTTATAGCTTTCTTATACCTGCTTTGGAAATCTTTGTCTATGGCATGAGA 1320
Db 1189 TTTGCTGAGGTTATAGCTTTCTTATACCTGCTTTGGAAATCTTTGTCTATGGCATGAGA 1248
Qy 1321 TCTTCATTTAAAGCTGAAATATCACTACGCTATGTCATCAAAATCCTTTGTGTGT 1380
Db 1249 TCTTCATTTAAAGCTGAAATATCACTACGCTATGTCATCAAAATCCTTTGTGTGT 1308
Qy 1381 GATTGCTGATGGGTTGTTACTTGTCTTTGTGGCATTTTGGATATAAAATACCGAGG 1440
Db 1309 GATTGCTGATGGGTTGTTACTTGTCTTTGTGGCATTTTGGATATAAAATACCGAGG 1368
Qy 1441 CAGTATCAAGAAATAGCTTGTCTGTGATGAGAGACGCTGACATGCCCTCATGGGGTTC 1500
Db 1369 CAGTATCAAGAAATAGCTTGTCTGTGATGAGAGACGCTGACATGCCCTCATGGGGTTC 1428
Qy 1501 CTGGCCATGCTGTCCACCGAAATCTCTGTCTGTACTGATCACTTGTGAGAGAG 1560
Db 1429 CTGGCCATGCTGTCCACCGAAATCTCTGTCTGTACTGATCACTTGTGAGAGAG 1488
Qy 1561 TTCTGCTGATGTTCTTCCCTTCAAGTAAATTCGACCTGAGAAACGCGACGCTCACTC 1620
Db 1489 TTCTGCTGATGTTCTTCCCTTCAAGTAAATTCGACCTGAGAAACGCGACGCTCACTC 1548
Qy 1621 ATCCCATTTGATCTGATGAGCGGATTTTAAATAGCTGTAATTCATTTTGGAAATAG 1680
Db 1549 ATCCCATTTGATCTGATGAGCGGATTTTAAATAGCTGTAATTCATTTTGGAAATAG 1608
Qy 1681 GATTATTTTGGAAACCTTTATGAGAAATATGAGATGTTTCCCATTTATATATACCA 1740
Db 1609 GATTATTTTGGAAACCTTTATGAGAAATATGAGATGTTTCCCATTTATATATACCA 1668
Qy 1741 ACAGAAATATTTGAGCAAGGATTTCTCTTGAATTTTCTTGAAGTGTGAACTTGTCTG 1800
Db 1669 ACAGAAATATTTGAGCAAGGATTTCTCTTGAATTTTCTTGAAGTGTGAACTTGTCTG 1728
Qy 1801 GCTTTTCTCATGTTGTTTCTTAAATTAATAGTCTCTGTCATTTCAAAAACCGCC 1860
Db 1729 GCTTTTCTCATGTTGTTTCTTAAATTAATAGTCTCTGTCATTTCAAAAACCGCC 1788
Qy 1861 TTGCAGACCAACAGATAGAAATGTTTGGAAAGAGTGGCTGTTGCAATGCTTTC 1920
Db 1789 TTGCAGACCAACAGATAGAAATGTTTGGAAAGAGTGGCTGTTGCAATGCTTTC 1848
Qy 1921 TTTTATATAGTGTCTCTGATGCCATCTGCTGATTTCTGTATTTGTGATTAATCTCT 1980
Db 1849 TTTTATATAGTGTCTCTGATGCCATCTGCTGATTTCTGTATTTGTGATTAATCTCT 1908
Qy 1981 TCCCTCTCCGGGATGAAATATACAGACAAATGACTCTCTGATAGTGAATTTTCTCT 2040
Db 1909 TCCCTCTCCGGGATGAAATATACAGACAAATGACTCTCTGATAGTGAATTTTCTCT 1968
Qy 2041 CCAGTTAAACAGTCTTTGATTCCTCTATCTCTCAACCAACTTTTATTAAGAC 2100
Db 1969 CCAGTTAAACAGTCTTTGATTCCTCTATCTCTCAACCAACTTTTATTAAGAC 2028
Qy 2101 AAGTTGAAACAGCTGCTGCAAAATCTCAGAGAAATCAATTTTCAAAATTAAGAA 2160
Db 2029 AAGTTGAAACAGCTGCTGCAAAATCTCAGAGAAATCAATTTTCAAAATTAAGAA 2088
Qy 2161 AAGTTATCTACATCTGTTGTGATAGAGACTCTCTCTGCTGAACTTGGGGTTTG 2220
Db 2089 AAGTTATCTACATCTGTTGTGATAGAGACTCTCTCTGCTGAACTTGGGGTTTG 2148

Qy 2221 AACAAATTAACCTTGAGAGACATATATAAAGAACAGTTTCC 2262
Db 2149 AACAAATTAACCTTGAGAGACATATATAAAGAACAGTTTCC 2190

RESULT 8
AX451566 2142 bp DNA linear PAT 03-JUL-2002
LOCUS AX451566
DEFINITION Sequence 5 from Patent WO0226824.
ACCESSION AX451566
VERSION AX451566.1 GI:21698550
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Feder, J.N., Muntler, G., Ramanathan, C.S. and Hawken, D.R.
A novel human g-protein coupled receptor, hgrbm5, expressed
highly in brain and ovarian tissues
Patent: WO 0226824-A 5 04-APR-2002;
Bristol-Myers Squibb Company (US)
Location/Qualifiers
1..2142

FEATURES
source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 613 a 435 c 408 g 686 t

ORIGIN

Query Match 90.9%; Score 2060; DB 6; Length 2142;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

Qy 52 ATGTTCTTTGACTTCAATTCATGCTGTGATCAATGTCAAAGATTGCACTGACTCA 111
Db 1 ATGTTCTTTGACTTCAATTCATGCTGTGATCAATGTCAAAGATTGCACTGACTCA 60
Qy 112 GGTGACATGATCAGCTCCCTTCAAGCAAGAAATTTTCCCTGAGGAACTTATACCAAG 171
Db 61 GGTGACATGATCAGCTCCCTTCAAGCAAGAAATTTTCCCTGAGGAACTTATACCAAG 120
Qy 172 TGCTTACCCCGAGCTTTTCACTGATGAGCAAGATGACTGTGGGACGGGCGGACGA 231
Db 121 TGCTTACCCCGAGCTTTTCACTGATGAGCAAGATGACTGTGGGACGGGCGGACGA 180
Qy 232 GAGAACTGTGTGACATGATGAGATGGGCGACATATTTGGCAACATGCAATGAAATGCT 291
Db 181 GAGAACTGTGTGACATGATGAGATGGGCGACATATTTGGCAACATGCAATGAAATGCT 240
Qy 292 AACGCGGAGCTTAAACACAGAGATGCTTTCTAAACAGATATCCAGATGCTGATGCTG 351
Db 241 AACGCGGAGCTTAAACACAGAGATGCTTTCTAAACAGATATCCAGATGCTGATGCTG 300
Qy 352 AAAGAACTGAATGGAATGTGAATGTGAATGCTTAAAGTCTGTGCGAGATTTCTAAC 411
Db 301 AAAGAACTGAATGGAATGTGAATGTGAATGCTTAAAGTCTGTGCGAGATTTCTAAC 360
Qy 412 AATGTGACATTAATCTGCTTTAAGAAAAACAAATCCACAGCTCTTCCAGATTAAGTTTC 471
Db 361 AATGTGACATTAATCTGCTTTAAGAAAAACAAATCCACAGCTCTTCCAGATTAAGTTTC 420
Qy 472 ATCAAAATACAAAACCTTAAAGATATTTCTTGACATTAATGATGATAGACATATCC 531
Db 421 ATCAAAATACAAAACCTTAAAGATATTTCTTGACATTAATGATGATAGACATATCC 480
Qy 532 AGAAAGATTTTGTGATATGTATCTGCAAAATATATATATCAACCAACTGATC 591
Db 481 AGAAAGATTTTGTGATATGTATCTGCAAAATAT----- 517
Qy 592 ACAACCTCAACCTGGAATATTCAAAGACTTAATCAATGCTTAATGCTTAATTTTAT 651

Db 39 CATGGATTAGTGTCAAAATGATGACAGATGATTCGATTTGACATGACTCAAGGTAG 98
Qy 117 CATGATACATCCTTCATGCGCAAAAAGATATTTCCCTGGGGAACTTTCACAAAGCTT 176
Db 99 CATGATACATCCTTCATGCGCAAAAAGATATTTCCCTGGGGAACTTTCACAAAGCTT 158
Qy 177 ACCCGAGCTTTTCACTGTATGCGCAAGATGATCTGTGGGACGGGGCGAGCGAAGAA 236
Db 159 ACCCGAGCTTTTCACTGTATGCGCAAGATGATCTGTGGGACGGGGCGAGCGAAGAA 218
Qy 237 CTGTGTGACACTATGATGATGGGCGACCATATTTGGGCACTGTGATGGAAAATGCTAACG 296
Db 219 CTGTGTGACACTATGATGATGGGCGACCATATTTGGGCACTGTGATGGAAAATGCTAACG 278
Qy 297 CGTGGCCCTTAACAGAGAGTCTTTCTAAACAGATTCACAAGCTGTGACTGCAAGA 356
Db 279 CGTGGCCCTTAACAGAGAGTCTTTCTAAACAGATTCACAAGCTGTGACTGCAAGA 338
Qy 357 AACTGAATGGATATGTAAATGATGACTTAAAGCTGTGCGGATGATTTCTAACATGT 416
Db 339 AACTGAATGGATATGTAAATGATGACTTAAAGCTGTGCGGATGATTTCTAACATGT 398
Qy 417 GACATTACTGTCTTAAAGAAAACAATTCACAGTCTTCAGATTAAGTTTTCATCA 476
Db 399 GACATTACTGTCTTAAAGAAAACAATTCACAGTCTTCAGATTAAGTTTTCATCA 458
Qy 477 ATACCAAACTTAAAGATATTTCTCAGATTAATGATTAATGACATATTCAGAGAA 536
Db 459 ATACCAAACTTAAAGATATTTCTCAGATTAATGATTAATGACATATTCAGAGAA 518
Qy 537 AGCATTTTGGATATGATTAATGCAAAATATTAATGCTCAACCAACCTGATCACAAC 596
Db 519 AGCATTTTGGATATGATTAATGCAAAATATTAATGCTCAACCAACCTGATCACAAC 578
Qy 597 CCTCAGACTGGAATATTCAAAGACTTACATCAGCTAACTGGCTTAATTCAGATGAGAA 656
Db 579 CCTCAGACTGGAATATTCAAAGACTTACATCAGCTAACTGGCTTAATTCAGATGAGAA 638
Qy 657 TCACATTAACAGAAATTCACAGCGCTGTGTTACGGGATTAATTCCTGTTTCTGTC 716
Db 639 TCACATTAACAGAAATTCACAGCGCTGTGTTACGGGATTAATTCCTGTTTCTGTC 698
Qy 717 TATGTTATTAATCTACTAGAACTCTTCCCAAGCAGATGTGTGCGCAATGCTCAACT 776
Db 699 TATGTTATTAATCTACTAGAACTCTTCCCAAGCAGATGTGTGCGCAATGCTCAACT 758
Qy 777 CAACCTGGGTGATTTGGAAAGCAATGAATTAAGTATCTCACAAATTCATGTTCTGTC 836
Db 759 CAACCTGGGTGATTTGGAAAGCAATGAATTAAGTATCTCACAAATTCATGTTCTGTC 818
Qy 837 GTGGATTCGCTCAAGTGTCTTCTGCTAGAAATCAAAATGTTTTGTTCCAGAGAA 896
Db 819 GTGGATTCGCTCAAGTGTCTTCTGCTAGAAATCAAAATGTTTTGTTCCAGAGAA 837
Qy 897 GACATTTTCTTCAATTAATAATTTAGAGAACTGATCTGTAGCAATACGATACGGA 956
Db 838 GACATTTTCTTCAATTAATAATTTAGAGAACTGATCTGTAGCAATACGATACGGA 866
Qy 957 GGTATACCTCAGCTTTTAAAGCTTGAAGCTTCTACAAAGCTGAACTGTCTATCCA 1016
Db 867 ACTATACCTCAGCTTTTAAAGCTTGAAGCTTCTACAAAGCTGAACTGTCTATCCA 926
Qy 1017 TCCCTCTATATCTTCAAGAAACAGATTTGAAAGCTTAAACAATCTCAGTCTAGA 1076
Db 927 TCCCTCTATATCTTCAAGAAACAGATTTGAAAGCTTAAACAATCTCAGTCTAGA 986
Qy 1077 CCTGGAAGAGATAGAGATTCAAATTAACAACAAGATGTTCAACCATAGAAATCT 1136
Db 987 CCTGGAAGAGATAGAGATTCAAATTAACAACAAGATGTTCAACCATAGAAATCT 1046
Qy 1137 TTCTCAAT-----TTATTTAAAAAATTTGATCTGTCTCTATGCTCCCATGTCCG 1190
TTCTCAAT-----TTATTTAAAAAATTTGATCTGTCTCTATGCTCCCATGTCCG 1190

Db 1047 TTCTCAATACCTGTATATTTCAAAAATTTGCAATACTGCTCTATGCTCCCATGTCCG 1106
Qy 1191 AATATGATATCCTTGAACGAGCAATTTCTCAATTTGAGACCTTGTGCTAACATAT 1250
Db 1107 AATATGATATCCTTGAACGAGCAATTTCTCAATTTGAGACCTTGTGCTAACATAT 1166
Qy 1251 CCTGAGAAATTTGTCTGGGTATATGCTTCAATTTACTGCTTTGGAAATCTTTTGTCTAT 1310
Db 1167 CCTGAGAAATTTGTCTGGGTATATGCTTCAATTTACTGCTTTGGAAATCTTTTGTCTAT 1226
Qy 1311 TGGCATAGATCTTTCAATTAAGCTGAAATTAACAATCAGAGCTATGCTCAATCAAAATCCT 1370
Db 1227 TGGCATAGATCTTTCAATTAAGCTGAAATTAACAATCAGAGCTATGCTCAATCAAAATCCT 1286
Qy 1371 TTGTGTGCTGATGCTGATGAGGTGTTTACTGTTCTTTTGTGGCAATTTTGCATATATA 1430
Db 1287 TTGTGTGCTGATGCTGATGAGGTGTTTACTGTTCTTTTGTGGCAATTTTGCATATATA 1346
Qy 1431 ATACCGAGGCAATATCAGAAATATGCTTGTGTGTGATGAGAGCTGCAAGTCCGCT 1490
Db 1347 ATACCGAGGCAATATCAGAAATATGCTTGTGTGTGATGAGAGCTGCAAGTCCGCT 1406
Qy 1491 CATGGGTTCTGGCCCATGCTGTCACCGAAGTCTGTTCTGCTACTGACCTTATGAC 1550
Db 1407 CATGGGTTCTGGCCCATGCTGTCACCGAAGTCTGTTCTGCTACTGACCTTATGAC 1466
Qy 1551 TTGGAAGAATTCCTGTGCTATTTGCTTCCCTTCAGTAACTTGCACCTGGAAAAAGCA 1610
Db 1467 TTGGAAGAATTCCTGTGCTATTTGCTTCCCTTCAGTAACTTGCACCTGGAAAAAGCA 1526
Qy 1611 GACCTCAGTATCTCATTTTGATCTGATGAGGCGGATTTTAAATAGCTGTAATTCAT 1670
Db 1527 GACCTCAGTATCTCATTTTGATCTGATGAGGCGGATTTTAAATAGCTGTAATTCAT 1586
Qy 1671 TTGGAATPAAGATTAATTTGGAAACTTTTAAAGGAAAAATGAGATATGTTCCCACTTA 1730
Db 1587 TTGGAATPAAGATTAATTTGGAAACTTTTAAAGGAAAAATGAGATATGTTCCCACTTA 1646
Qy 1731 TTATGACCAACAGAAATATTTGAAACAAAGGATATCTCTTGAAATTTTCTAGGTGT 1790
Db 1647 TTATGACCAACAGAAATATTTGAAACAAAGGATATCTCTTGAAATTTTCTAGGTGT 1706
Qy 1791 GAACCTGTGCTTTTCTCATATGTTGTTTCTATATTAATTAATGTTCTGATTTCA 1850
Db 1707 GAACCTGTGCTTTTCTCATATGTTGTTTCTATATTAATTAATGTTCTGATTTCA 1766
Qy 1851 AAAAAGCCCTTGAAGACCAAGAAATGAGAAATTTGTTGAAAGAGAGTGGCTGTGC 1910
Db 1767 AAAAAGCCCTTGAAGACCAAGAAATGAGAAATTTGTTGAAAGAGAGTGGCTGTGC 1826
Qy 1911 AAATCGTTCTTTTATATGTTTCTCTGATGCAATCTGCTGATTTCTGATATTTGAT 1970
Db 1827 AAATCGTTCTTTTATATGTTTCTCTGATGCAATCTGCTGATTTCTGATATTTGAT 1886
Qy 1971 TAAATCTTTTCCCTTCTGCGGGTGAATATCCAGACCAATGACTTCTGATATTTGAT 2030
Db 1887 TAAATCTTTTCCCTTCTGCGGGTGAATATCCAGACCAATGACTTCTGATATTTGAT 1946
Qy 2031 TTTTTCCTTCAGTTACAGTGTGTTGAATTCATCTCTATATCTCTCAACAACCACT 2090
Db 1947 TTTTTCCTTCAGTTACAGTGTGTTGAATTCATCTCTATATCTCTCAACAACCACT 2006
Qy 2091 TTTTAAAGCAAGTTGAACAGTGTGCTGCAAAACATCAGAGAAATCAATTTTCAAT 2150
Db 2007 TTTTAAAGCAAGTTGAACAGTGTGCTGCAAAACATCAGAGAAATCAATTTTCAAT 2066
Qy 2151 TTTTAAAGCAAGTTGAACAGTGTGCTGCAAAACATCAGAGAAATCAATTTTCAAT 2210
Db 2067 TTTTAAAGCAAGTTGAACAGTGTGCTGCAAAACATCAGAGAAATCAATTTTCAAT 2126
Qy 2211 TGGGTTTGAACAATAATACATCTGAGAGATTAATGAACCAAGTTTCTAG 2265
Db 2127 TGGGTTTGAACAATAATACATCTGAGAGATTAATGAACCAAGTTTCTAG 2181

RESULT 10
 AX385037
 LOCUS AX385037 2049 bp DNA linear PAT 19-MAR-2002
 DEFINITION Sequence 11 from Patent WO0214489.
 ACCESSION AX385037
 VERSION AX385037.1 GI:19578160
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
 TITLE Leucine-rich repeat-containing g-protein coupled receptor-8
 JOURNAL Patent: WO 0214489-A 11 21-FEB-2002;
 Amgen, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..2049
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 1..2049
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28590.1"
 /db_xref="GI:19578161"
 /translation="MIVPLVFKFLFSLRLITMPLFLHIVLINVKDFALTOGSMITPS
 COKGYPFGCNLTCLPRAFHCDGDKDGNADENCGDPSGMATIGYVHGNANVAL
 TOECFLKOYBOCCDCKETELCEVNGDLKSPMSNNVTLSLTKNTHSLPKVPEIKY
 TKKKIPLQNCICRHSRKAPFSLCQLIILIDNPITRISQPLFGNLSLFLSMVN
 NVLEALPKQACQMPQLNWDLEGNRIKIVNSTFLSCSLVLIDLSNITELSHL
 FKDLKLOKDLERIBIPNINRMFQPMKLSHIYKPNRYCSYAPHRVIMPELTGII
 SSBEDLANNILRIFWVIAFTFCFGLPIYKMSFIKENTHMSIKILCCADCLM
 GYVLPFVGIYFDIKYGOYOKYLLMESVOCRLMGFLAMSTEVSLILTYLLEKEI
 VYFPPSNIRPGKORTSVILICIMWAGFLIAVLPNNKQYFGNPKNGVCPFLVYD
 TEDIQKYSLSGIPLGVNLLAFIIVPSYITMESCIOKRLQOTTEPRANCGPVAANO
 RPFPIVSDAICWIPVAVKISLFRVEIDPTNLSIVIFLPSNALSALPILTLTIN
 FFDKLLKQLLHKRQKSIPIKIKKSLSTISWIEDSSSLKGLVINKITGLDSIMKVS
 "

sig_peptide 1..108
 BASE COUNT 579 a 415 c 395 g 660 t
 ORIGIN

Query Match 79.6%; Score 1803; DB 6; Length 2049;
 Best Local Similarity 90.5%; Pred. No. 0;
 Matches 2049; Conservative 0; Mismatches 0; Indels 216; Gaps 3;

QY 1 ATGATGTTTTCGGTTTAAACATCTTCAGCTCAGATTGATTACATGTTCTT 60
 DB 1 ATGATGTTTTCGGTTTAAACATCTTCAGCTCAGATTGATTACATGTTCTT 60
 QY 61 CTAATGATTCATCGTTTGATCATGTCAGAAATTTGATGATCAAGTAGCATG 120
 DB 61 CTAATGATTCATCGTTTGATCATGTCAGAAATTTGATGATCAAGTAGCATG 120
 QY 121 ATCACTCTTCATGATGCGGAGATATTTTCCCTGTGGGAATCTTACCAAGTCTTACC 180
 DB 121 ATCACTCTTCATGATGCGGAGATATTTTCCCTGTGGGAATCTTACCAAGTCTTACC 180
 QY 181 CGAGCTTTTCACTGTATGGCAAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 240
 DB 181 CGAGCTTTTCACTGTATGGCAAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 240
 QY 241 GGTGACACTAGTATGGGCGACCATATTGTCAGAGTGAATGCTAACAGCGGTG 300
 DB 241 GGTGACACTAGTATGGGCGACCATATTGTCAGAGTGAATGCTAACAGCGGTG 300
 QY 301 GCCTTAACACAGAGAGTCTTTCTAAGAACATACCAATCTGCTGCAAGAAACT 360
 DB 301 GCCTTAACACAGAGAGTCTTTCTAAGAACATACCAATCTGCTGCAAGAAACT 360

QY 361 GAATGGAATGTGTAAATGAGTGAATTAAGTCTGTGCGGATGATTTCTAACATGTGACA 420
 DB 361 GAATGGAATGTGTAAATGAGTGAATTAAGTCTGTGCGGATGATTTCTAACATGTGACA 420
 QY 421 TTAATGCTCTTAAAGAAACAAATATCCAGATCTTCCAGATTAAGTTTCTCAAAATAC 480
 DB 421 TTAATGCTCTTAAAGAAACAAATATCCAGATCTTCCAGATTAAGTTTCTCAAAATAC 480
 QY 481 ACAAAACCTTAAGAAATATTTCTCAGCATTAATGATGATGACATATCCAGAAAGCA 540
 DB 481 ACAAAACCTTAAGAAATATTTCTCAGCATTAATGATGATGACATATCCAGAAAGCA 540
 QY 541 TTTTGTGATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 DB 541 TTTTGTGATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 QY 601 AGACCTGGAATATTAAGAACTTACATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 601 AGACCTGGAATATTAAGAACTTACATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 661 ATTAACGAAATTTCAACAGCTTGTGTTTACGGAATTAATTCCTGTTTCTGCTATG 720
 DB 661 ATTAACGAAATTTCAACAGCTTGTGTTTACGGAATTAATTCCTGTTTCTGCTATG 720
 QY 721 GTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 DB 721 GTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 QY 781 TGGGTGATTTGGAAGCAATGAATTAAGTATCTCAAAATTTCAAGTTTCTGCTGTC 840
 DB 781 TGGGTGATTTGGAAGCAATGAATTAAGTATCTCAAAATTTCAAGTTTCTGCTGTC 840
 QY 841 GATTGCTCAACGCTGCTTCTGCTGCAAAATCAAAATGTTTGTTCAGAGAAAGACA 900
 DB 841 GATTGCTCAACGCTGCTTCTGCTGCAAAATCAAAATGTTTGTTCAGAGAAAGACA 900
 QY 901 TTTTCTTCAATTAAGAAATTTAGAGAACTGATCTGTAGAAATGATGATGATGATGATGAT 960
 DB 901 TTTTCTTCAATTAAGAAATTTAGAGAACTGATCTGTAGAAATGATGATGATGATGATGAT 960
 QY 961 TCACCTCACTTTTAAAGACTTGAAGCTTCTTCAAAAGCTGAACCTGATCAATCTCT 1020
 DB 961 TCACCTCACTTTTAAAGACTTGAAGCTTCTTCAAAAGCTGAACCTGATCAATCTCT 1020
 QY 1021 CTTATGATCTTCAAGAAACCAAGTTGAAGATTTAAACATTCAGTCTGAGACCTG 1080
 DB 1021 CTTATGATCTTCAAGAAACCAAGTTGAAGATTTAAACATTCAGTCTGAGACCTG 1080
 QY 1081 GAAAGATAGAGATTCGAATATTAACACAGAAATGTTTCAACCCATGAAGATCTTCT 1140
 DB 1081 GAAAGATAGAGATTCGAATATTAACACAGAAATGTTTCAACCCATGAAGATCTTCT 1140
 QY 1141 CACATTTATTTCAAAACCTTGGATATGCTGCTCTATGCTCCCATGTCGAATATGATG 1200
 DB 1141 CACATTTATTTCAAAACCTTGGATATGCTGCTCTATGCTCCCATGTCGAATATGATG 1200
 QY 1201 CCGTTAGCGAGCACTTCTTCAATTTGAGACCTTGTGGCTAAAGATATCCAGAAAT 1260
 DB 1201 CCGTTAGCGAGCACTTCTTCAATTTGAGACCTTGTGGCTAAAGATATCCAGAAAT 1260
 QY 1261 TTTGTCTGGTTTATGATCTTCAATTCATGCTTGTGAATCTTTTGTGATGATGATGATGAT 1320
 DB 1261 TTTGTCTGGTTTATGATCTTCAATTCATGCTTGTGAATCTTTTGTGATGATGATGATGAT 1320
 QY 1321 TCTTTTCAATTAAGCTGAATATCAATCACTCACTATGCTCAATTAATCTTTGTGTGCT 1380
 DB 1321 TCTTTTCAATTAAGCTGAATATCAATCACTCACTATGCTCAATTAATCTTTGTGTGCT 1380
 QY 1381 GATTGCTGATGAGGATTTTACTGTTCTTTGTTGGCAATTTTGCATATTAATTAATCCAGAGG 1440
 DB 1381 GATTGCTGATGAGGATTTTACTGTTCTTTGTTGGCAATTTTGCATATTAATTAATCCAGAGG 1440
 QY 1440 GATTGCTGATGAGGATTTTACTGTTCTTTGTTGGCAATTTTGCATATTAATTAATCCAGAGG 1500
 DB 1440 GATTGCTGATGAGGATTTTACTGTTCTTTGTTGGCAATTTTGCATATTAATTAATCCAGAGG 1500

[illegible]

JOURNAL	Genes 30 (1), 26–35 (2001)
MEDLINE	21250990
PUBMED	11353515
REFERENCE	2 (bases 1 to 2539)
AUTHORS	Agoulnik, A.I.
TITLE	Submitted (06-FEB-2001) Ob/Gyn, Baylor College of Medicine, 6550 Famin Str. Su. 861, Houston, TX 77030, USA
JOURNAL	Submitted (06-FEB-2001) Ob/Gyn, Baylor College of Medicine, 6550 Famin Str. Su. 861, Houston, TX 77030, USA
FEATURES	
Source	Location/Qualifiers
gene	1..2539
CDS	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6"
	/db_xref="taxon:10090"
	/chromosome="5"
	1..2539
	/gene="Great"
	194..2407
	/gene="great"
	/note="GREAT"
	/codon_start=1
	/product="G protein coupled receptor affecting testicular descent"
	/protein_id="AA08943.1"
	/db_xref="GI:15811373"
	/translation="MTMLIHVTLTEVDPALADSSMVA.PI.CPKGYPFGCN.LTKC.PI.APHCDGVDDCGNAGEDNCGDTSWTTT.PGYTHGVNKKYTLLOBCTLSOYPOHCYR.NELBEKVADLAVPRVSSNVTLSLKKNKIL.PVKSRYELRKIVLOHNTHTIRRA.FLGNILQILVLSHNCITSLRPGIFEDLHQLAML.IIDNPITRI.SOKSEMGNSL.FEIPVGNRLEAL.PETTLCAQMPQLMNVDLANNIGIKYITNSTPCLSLTVL.FLPRNQ.GFVPEKTPSSLNKLGEIDLSNMITKLPVLPSPDLHLOKMLS.SNPDLVHNQOQ.LKOLISDLERI.EI.PNISTGMFQPKONLSHIVLKTIPRYCSYPHVR.ICMPTSDIGIS.EDILANGILRYSVWVIAFRTTCGPNLIVAVBLIKABNTYAMS.IKILCADCLMNV.LPSGVDPDKTQYOKYAL.MBESVPCKLDFLTLTSTVSLTLFPLTEFLYIV.FPSSRLYKGRQYALVALS.IWVGF.LI.AVPTREIDY.FGNFYKGVGVCPLHVDQAE.FGSRYSLSIGIFLGNVLA.FLVIVISYVMFCS.IHKTALQTAVRSHI.GEYVAVNRF.FIVSDACIMV.IPVFVKILSLLOVEI.PGTTISMI.VVFLPVNSALNP.IIYVLTTSF.FDKLQILHKKRRKPI.FFKVKKSLSI.SI.VWTDES.LKGLVLSKIALGDS.IMKPVS"
BASE COUNT	639 a 641 c 574 g 685 t
ORIGIN	
Query Match	69.7%; Score 1579.8; DB 10; Length 2539;
Best Local Similarity	82.1%; Pred. No. 0;
Matches 1831; Conservative	0; Mismatches 397; Indels 3; Gaps 1
32	TCAGCCTCAGATTGATTACATGTTCTTCTACCTTCATTCACTCGTTCGATCAAGTCA 91
Db	174 TCGAGCGACACCGCGCCGCAATGTGGCTCTCTACTTCATGTCATCCTTCTCGACAGAGTCA 233
92	AAGATTTGGCACTGACTCAAGGTAGCATGATACCTCCCTCAGGCCCAAAAAGATATTTC 151
Db	234 AAGATTTTGGCACTGCTGCTGACACGATTAAGTGGCTCTCTGTCGCCCAAGGTAATTTTC 233
152	CCTGTGGAAATCTTCAACAAAGTCTTACCCCGAGCTTTTCACTGTGATGCGCAAGATGACT 211
Db	294 CCTGTGGAAATCTTCAACAAATGCTTGTGCCCGAGCTTTTCACTGTGATGCGTGTGATGAT 353
212	GTGGGAACGGGGCGGACCGAAGGAACCTGTGTGTGACACTAGTGGATGGCGGACCATATTTC 271
Db	354 GCGGGAATGGTCCGACGACGAGGCAACCTGTGTGTGACACTAGTGGATGGCAACCATATTTC 413
272	GCAACGTGATGGAATGCTTAAACAGCGTGGCCTTAAACAGAGAGTGTCTTTTAAACAGT 331
Db	414 GCAACGTGATGGAATGCTTAAACAGTGAATTAAGTGAACATTGACACAGAGTGTCTTTTCAAGCAGT 473
332	ATCCACAAATGCTGTGACTGTCGAAAGAACTGAAATGGAAATGTGTAATGCTGACTTAAAGT 391
Db	474 ATCCACACACACTGTACTGTCGAGAGAAATGAATCGGAATGTGTAAAGCTGACTTAAAG 533
392	CTGTGCCGATGATTTCTTAAACAATGAGCACTTACGTCTCTTAAAGAAAACAAATCCACA 451
Db	534 CTGTGCCAAGGTCTTCAAGCAACGTAACCTTACATCTCTTAAAGAAAACAAATCCACA 593

QY 452 GTCTCCAGTAAAGTTTCATCAATACCAAACTTAAAGATATTTCTCAGCAT 511
 DB 594 GACTTCAGCAAGGCTTCAGAGATACAGAAAGATATACCTTCAGACAC 653
 QY 512 ATTCATTTAGACATATCCAGAGAAAGCATTTTGGATTATGTAACTTCAATATAT 571
 DB 654 ACTGCATACACATCTCAGAGAGACATTCCTTGATTAATATCTCAAAATCTGT 713
 QY 572 ATTCACCAACAACCTGATCAACCTCAGACCTGGAATATTCAGAACTTACATCAG 631
 DB 714 ATTCAGCCCAATATGATATACCTCTCAGGCTGGGATATTCAGAACTTCATCAGC 773
 QY 632 TAACTGGCTAAATTCATGATGACATCAATCAATTCAGAGGCTGTGTTACGG 691
 DB 774 TTGCTTGCTAAATTTAGATGACACCCGATCACCAATCTCAGAAAGCTCTTATGG 833
 QY 692 GATTAATTCCTTGTCTGTCTGTATGTTATTAATCTACTTGAAGCTTTCCAGC 751
 DB 834 GGTAAACCTCTGTCTGTCTGTCTGTATGTTATTAATCTACTTGAAGCTTTCCAGC 893
 QY 752 AGATGTGCTCCAAATGCTCACTCACTGGGTGATTTGGAAGCAATGAATTAAGT 811
 DB 894 CATGTGTGCTCAGATGCTCACTCACTGGGTGATTTGGAAGCAATGAATTAAGT 953
 QY 812 ATTCACAAATTCAGCTTCTGTCTGTATGTTATTAATCTACTTGAAGCTTTCCAGC 871
 DB 954 ACATTAAGAACTCAGCTTCTGTCTGTATGTTATTAATCTACTTGAAGCTTTCCAGC 1013
 QY 872 ATCAATTTGTTTGTTCAGAGAGACATTTCTTCAATTAATTAAGAGAACTGG 931
 DB 1014 ATCAATTTGTTTGTTCAGAGAGACATTTCTTCAATTAATTAAGAGAACTGG 1073
 QY 932 ATCTGTATGATATGATTAAGAGAGATCACTCACTTTTAAAGCTTGAAGCTTC 991
 DB 1074 ACCGTCTAGCAATATATTAACAAACTCCAGTCCACTTTTGAAGCACTTCATCTTC 1133
 QY 992 TACAAAGCTGAACCTGTCACTCACTTTATGATATCTTCAATTAATTAAGAGAACTGG 1051
 DB 1134 TCCAGAAAGCTGAACCTGTCACTCACTTTATGATATCTTCAATTAATTAAGAGAACTGG 1193
 QY 1052 GTCTTAAACAATTCAGCTTCTGTATGTTATTAATCTACTTGAAGCTTTCCAGC 1111
 DB 1194 GTCTCAAAACAATTCAGCTTCTGTATGTTATTAATCTACTTGAAGCTTTCCAGC 1253
 QY 1112 GAATGTTCAACCATGAAGATCTTTCTCAATTAATTAAGAGAACTTCATGATCTGT 1171
 DB 1254 GAATGTTCAACCATGAAGATCTTTCTCAATTAATTAAGAGAACTTCATGATCTGT 1313
 QY 1172 CCTATGCTCCCATGTCGAAATATGATGCTTGAAGAGAGAGATTTCTTCAATTAAG 1231
 DB 1314 CCTATGCTCCCATGTCGAAATATGATGCTTGAAGAGAGAGATTTCTTCAATTAAG 1373
 QY 1232 ACCCTTGGCTAACAATTCCTCAGAAATATTTGTCTGGGTATTAAGCTTTCAATTAAG 1291
 DB 1374 ACCCTTGGCTAACAATTCCTCAGAAATATTTGTCTGGGTATTAAGCTTTCAATTAAG 1433
 QY 1292 TTGGAATCTTTTGTCAATGAGATCTTTCAATTAAGAGAGAGATTTCTTCAATTAAG 1351
 DB 1434 TTGGAATCTTTTGTCAATGAGATCTTTCAATTAAGAGAGAGATTTCTTCAATTAAG 1493
 QY 1352 CTATGTCATCAAAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1411
 DB 1494 CTATGTCATCAAAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1553
 QY 1412 TTGGCAATTTTGAATTAATTAACGAGAGAGATCAGAGATATGCTTGTGTGTGTGTGT 1471
 DB 1554 TTGGCAATTTTGAATTAATTAACGAGAGAGATCAGAGATATGCTTGTGTGTGTGTGT 1613
 QY 1472 AGAGGTGCTGAG 1531
 DB 1614 AGAGGTGCTGAG 1673
 QY 1532 TGCTATGACCTACTTGAATTTGAGAAAGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1591

DB 1674 TGCTGTGCTCAATTCCTGAG 1733
 QY 1592 TTGCACTTGAGAAACGAG 1651
 DB 1734 TGCGCTGAG 1793
 QY 1652 TAAATGCTGTAATTCATTTTGAATTAAGAGATTTATTTGAGAACTTTATGAGAAAG 1711
 DB 1794 TCAATGAG 1853
 QY 1712 GAGATGTTTCCCACTTTATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
 DB 1854 GAGTCTGCTCCCACTTATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1913
 QY 1772 TTGGAATTTTCTAG 1831
 DB 1914 TTGGGATTTTCTAG 1973
 QY 1832 CTATGTTTCTGCTCAATTAATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1891
 DB 1974 CCAATGTTCTGCTCAATTAATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2033
 QY 2034 GGAAG 2093
 DB 1952 GGAATTCCTGATTTGATTAATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2011
 QY 2094 GGAATTCCTGATTTGATTAATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2153
 DB 2012 TGACCTTCGAG 2071
 QY 2154 TCACTTCGAG 2213
 DB 2072 ATACTCTCAACAACCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2131
 QY 2214 ACACTGAG 2273
 DB 2132 GGAATTCATTTTCAATTAATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2191
 QY 2274 GGAATTCATTTTCAATTAATTAATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2233
 DB 2192 ACTCTCTTCCGAG 2251
 QY 2234 A---GTTTACTTAACTTGAG 2290
 DB 2252 AACCAAGTTTCC 2262
 DB 2291 AGCCGCTCTCC 2401

RESULT 12
 AX385045
 LOCUS 2214 bp DNA linear PAT 19-MAR-2002
 DEFINITION Sequence 19 from Patent WO0214489.
 ACCESSION AX385045
 VERSION AX385045.1 GI:19578166
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCES
 1. Paezzy, C.J., Gong, J., Daugherty, B. and Rogers, N.
 Leucine-rich repeat-containing g-protein coupled receptor-8
 molecules and use thereof
 Patent: WO 0214489-A 19 21-FEB-2002;
 Amgen, Inc. (US)
 FEATURES
 source 1..2214
 /organism="Mus musculus"
 /mol_type="genomic DNA"

CDS

/db_xref="taxon:10090"
1..2214
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28593.1"
/db_xref="gi:19578167"
/translation="MWLLHVLITLVKFDALADSSMYAFLPKGYFPCGNITKCLPR
AFHCDGVDCGADENCGDTSQWTTTIGTVGNVKNKTLQCEFLQYPOHCYRE
NEBECVADLKAIVKPVSNVTLSTKKKHRLPVKFSRYELKILQHNCHLHIS
RRALGLHNLILYSHNCITLSLRPIKFDLQHLAMLLDNDPIRISOKSPMGINSL
FPLSMVGRLELPETLCAQNPQLMVDLANNGIKYITNSTFLTCDSTLTVLPLPNOI
GFPEKITSLEKLEBLSNMITKLPHVLPFDHLQKLNLSNPILVYHKQFGS
LKQSLDLERLEINISGMQPMKNLSHLYLKTFRCSYVPHYRICMPSDQISS
EDLLAMGLIRSVWVIAFTCVGNFLVAVSLIKAEHTVAMSIKILCQDCLMGVY
LFSVGFPIKYRGQYKXALLMMSVPCRLGFLATLSTEVSLITLLEKFLVIV
FPSNLRLGKQTLAVALSIMWGFLLAIVPPTREDYDGNFYGKRGVCPPLHYDAED
FGSRGYSIGTILGVLLAFLVIVISYVTMFCGSHHTAQLQTAVERSHIKAVANRPF
FIVSDAICWLPFVIVKLISLQVEIPGITSIVVFLPVNSALNPILVUTTSFFK
DKLKQLHHRKRPFFVKKSLASIVWTBESSLKLQVLSKIALGDSIMKRVSP"

sig_peptide 554 a 548 c 492 g 620 t
BASE COUNT
ORIGIN

Query Match 69.7%; Score 1579; DB 6; Length 2214;
Beet Local Similarity 82.5%; Pred. No. 0;
Matches 1823; Conservative 0; Mismatches 385; Indels 3; Gaps 1;

52 ATGTCCTTCTACTCATTTGATCGTGTGATCAATGTCAAAGATTTTGACATGCA 111
1 ATGTGGCTCTTCACTTCACTCTCTTCTGACAGAGCTCAAGATTTTGACATGTCAG 60
112 GGTACATGATCACTCTTCTGATGCCAAAGATATTTTCCGTGGGAACTTACCAAG 171
61 AGCAGTATGTGGCTCTGTGCCCAAAGGTAATTTTCCGTGGGAACTTACCAAA 120
112 TGCTTAACCCGAGCTTTTCACTGTGACAGATGACTGTGGGAAAGGGGCGACAA 231
122 TGCTTAACCCGAGCTTTTCACTGTGACAGATGACTGTGGGAAAGGGGCGACAA 231
121 TGCTTGGCCGAGCTTTTCACTGTGACAGATGACTGTGGGAAAGGGGCGACAA 180
232 GAGAACTGTGTGACATGATGATGGGCGACATATTTGGACAGTGAAGAAATGCT 291
181 GACAACTGTGTGACATGATGATGACAACTATTTGGACAGTGCATGGGAATGTC 240
232 AACAGGTGGCTTAAACAGAGATGCTTTTAAACAGTATCCAAATGCTGTGACTGC 351
241 AATTAAGTGAATGACAGAGATGCTTTTCAAGCAGTATCCAAAGCTGTGATCTGC 300
352 AAAAGAACTGAATGGAATGTGTAATGTGACTTAAAGTCTGTGCCGATGATTTCTAAC 411
301 AGAGAAATGAAGCTGGAATGTGTAAGGCTGACTTAAAGCTGTGCAAAAGGTTTCCAGC 360
412 AATGTGACATTAAGTCTCTTAAAGAAACAAATATCCAGATCTTCCAGATTAAGTTTC 471
361 AACGTAACTTAATCTCTTAAAGAAACAAATATCCAGATCTTCCAGATTAAGTTTC 420
472 ATCAATACACAAACTTAAAGATATTTTTCAGCATTAATGCACTTAAGACATATCC 531
421 AGCAGATACACAAACTTAAAGATATTTTTCAGCATTAATGCACTTAAGACATATCC 480
532 AGGAAAGCAATTTTGGATTATGTAATGCAAAATTTATATCTCAACACATGCAATC 591
481 AGGAGGCAATCTTGGATTATGTAATGCAAAATCTGATCTCAGCATTAAGTCAAT 540
592 ACAACCTTCAGACTGGAATATTTCAAGACTTACATCAGCTAACTGGCTAATTTAGAT 651
541 AACTCTCTCAGGCTGGATATTTCAAGACTTGCACACTTGGCTAATTTAGAT 600
652 GACATTCACATTAACAGAAATTTCAAGCTGTGTTTACGAGATTAATTTCTTGTTC 711
601 GACAACTCCATACAGAAATTTCAAGAGTCTTATAGGGGTAAACTCTTGTTC 660
712 CTGTATAGGTTAATTAATTAAGAGCTCTTCCAAAGATGATGATCCAAAGCT 771

Db 661 TTGTTCATGATGGGTAACCGGCTCGAGGCGCTTCCCTGAACATGTGTGCTAGATGCT 720
Qy 772 CAATCAACCTGGGTGGAATTTTGAAGGCAATGAATTAAGTATTCACAAATTTCTACGTT 831
Db 721 CAATCAACCTGGGTGGAATTTTGAAGGCAATGAATTAAGTATTCACAAATTTCTACGTT 780
Qy 832 CTGTGTGCGATTCGCTCACAGTGTGTTTCTGCTAGTAAATTCAAATTTGTTTGTCCA 891
Db 781 CTAACTGCGACTGCTCACAGTGTGTTTCTGCTAGTAAATTCAAATTTGTTTGTCCA 840
Qy 892 GAGAAACATTTTCTTCACTTAATAAATTTAGAGAACTGATCTGTCTAGCAATACGATA 951
Db 841 GAGAAACATTTTCTTCACTTAATAAATTTAGAGAACTGATCTGTCTAGCAATACGATA 900
Qy 952 ACGAGCATATACCTCACCTTTTAAAGACTTGAAGCTTCAAAAGCTGAACCTGCA 1011
Db 901 ACAAACCTCCAGTCCACCTTTTCAAGCACTTCACTTCCAGAACCTGAACCTGCA 960
Qy 1012 TCCAACTCTTATGATCTTCAAGAAACGAGTTGAAAGCTTAAACAACTTCACTCT 1071
Db 961 TCCAACTCTTCTGTATGTCCACAAAGAACAGTTTGAAGTCTCAAACTTCACTCT 1020
Qy 1072 CTAGACTGGAAGATAGATTCCAATATTAACACAGATGTTCAACCCATGAAG 1131
Db 1021 CTAGACTGGAAGATAGATTCCAATATTAACACAGATGTTTCCAGCAATGAAG 1080
Qy 1132 AATCTTCTCATTTATTTAAACAACTTGAATGATGCTGCTGCTTAACTCCCATGCGGA 1191
Db 1081 AATCTTCTCATTTATTTAAACAACTTGAATGATGCTGCTGCTTAACTCCCATGCGGA 1140
Qy 1192 ATATGATGCCCTTGACGGAAGGCAATTTCTTCAATTTGAGAACCTTGGCTAAATATC 1251
Db 1141 ATCTGATGCCGCTGACATGATGATTTCTTCTGCTGAGAACCTTGGCTAAATATC 1200
Qy 1252 CTCAGAAATTTGTGCTGATTAAGCTTCAATACCTGCTTGAATCTTTTGTCTATT 1311
Db 1201 CTCAGAGTGTGTGCTGATTAAGCTTCAATACCTGCTTGAATCTTTTGTCTATT 1260
Qy 1312 GGCATGATCTTTCATTAAGCTGAAATCAATCACTGAGTATGCTCAATCAATCTT 1371
Db 1261 GCCGATGATCTTTCATTAAGCTGAAATCAATCACTGAGTATGCTCAATCAATCTT 1320
Qy 1372 TGTGTGCTGATTTGCTGATGAGGTGTTACTTGTCTTGTGGAATTTTGATTAATA 1431
Db 1321 TGTGTGCTGATTTGCTGATGAGGTGTTACTTGTCTTGTGGAATTTTGATTAATA 1380
Qy 1432 TACCGAGGCGATACAGAAATGATGCTTGTGCTGATGAGAGGCTGACAGTCCGCTTC 1491
Db 1381 TACCGAGGCGATACAGAAATGATGCTGCTGATGAGAGGCTGACAGTCCGCTTC 1440
Qy 1492 ATGGGCTTCGAGCATGCTGCCAGAGCTCTGTTCTGCTAGTACCTTGAATCT 1551
Db 1441 CTGGCTTCCTGAGCATGCTGCCAGAGCTCTGTTCTGCTAGTACCTTGAATCT 1500
Qy 1552 TTGAGAAATTCCTGATGATGCTTCCCTTCAAGTAACTTGAATCTTGAATCTTGA 1611
Db 1501 CTGAGAAATTCCTGATGATGCTTCCCTTCAAGTAACTTGAATCTTGAATCTTGA 1560
Qy 1612 AACTGATCACTTCAATTTGATGATGAGGGAATTTTAAATGCTGTAATTCATTT 1671
Db 1561 AACTGATCACTTCAATTTGATGATGAGGGAATTTTAAATGCTGTAATTCATTT 1620
Qy 1672 TGAATTAAGATTAATTTGAAACTTTATGGAATAATGAGATGTTTCCACTTAT 1731
Db 1621 ACCAGAGAGATTAATTTGGAACCTTTATAGGAATAATGAGATGTTTCCACTTAT 1680
Qy 1732 TATGACCAAGCAAGATTTGGAAGCAAGGATATTTCTTGGAAATTTTCTAGATG 1791
Db 1681 TATGACCAAGCAAGATTTGGAAGCAAGGATATTTCTTGGAAATTTTCTAGATG 1740
Qy 1792 AACTGCTGGCTTCTGATGATGATGTTTCTTAATTAATGATGCTGCTTCAAT 1851
Db 1741 AACTGCTGGCTTCTGATGATGATGTTTCTTAATTAATGATGCTGCTTCAAT 1800

Qy	1852	AAAAACCCCTTGGACAGACCA	CAGAAGTAAGAAATGTTTGGAGAGAGGTGGCTGTGGCA	1911
Db	1801	AAAACAGCCCTTACAGATCGACAGAAAGTAGAGAGCCACATCCGGGAAGAGAGTGGCTGTGGCA		1860
Qy	1912	AATCGTTTCCTTTTAAATAGTGTCTCTGATGCCATCTGCTGGAATTCCTGTATTTTGAGTT		1971
Qy	1972	AAATTCCTTCCCTCTCCGGGTGGAAATACAGACCAATGACACTCTCGATFAGTATT		2031
Db	1921	AAGATCCGTCTCTCCCTTCAAGTAGAGTACAGACCAATGACACTCTCGATFAGTATT		1980
Qy	2032	TTTTTCCCTCCAGTTAACAGTGTCTTGAATCCATTCCTCTTACTCTCACACCAACTTT		2091
Db	1981	TTTTTCCCTCCGGTGAACAGGCGCTTAAACCCATTCCTCTACACTCTGAGACCTCCCTTT		2040
Qy	2092	TTTAAAGCAAGTTGAAAACAGCTGCTGCACAAACATCAGAGAGAAATTCATTTTCAAAATT		2151
Db	2041	TTTAAAGCAAGTTGAAAACAGTGTCTGCACAAACATCAGAGAGAAACCATCTTCAAAAGTT		2100
Qy	2152	AAAAAAAAAAGTTATATACATTCATGTTGTGGAATGAGAGACTCTCTCCCTGAAACTT		2211
Db	2101	AAGAAGAAAATTTATCCGCTTCCATGTTGTGACAGACGA--GTCCTTCACTTAAACTT		2157
Qy	2212	GGGGTTTGAACAAAATATACACTTGGAGACAGTATTAATGAACAAGTTTCC		2262
Db	2158	GGAATGTTGACCAAAATAGCCCTTGGGACAGTATTAATGAAGCCGGTGTCC		2208
RESULT 13				
BD141175				
LOCUS	BD141175	1545 bp	DNA	linear
ACCESSION	BD141175			PAT 18-SEP-2002
VERSION	BD141175.1	GI:323236120		
KEYWORDS	WO 0204640-A/4.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 1545)			
JOURNAL	Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.			
	Novel G protein-coupled receptor protein and its DNA			
	Patent: WO 0204640-A 4 17-JAN-2002;			
	TAKEDA CHEMICAL INDUSTRIES LTD,TAKEO MORIYA,TAKASHI ITO, YASUSHI			
	SHINTANI, NOBUYUKI MIYAJIMA			
COMMENT	OS Homo sapiens (human)			
	PN WO 0204640-A/4			
	PD 17-JAN-2002			
	PF 06-JUL-2001 WO 2001JP005878			
	PR 07-JUL-2000 JP 00P 211989 18-DEC-2000 JP 00P 383794 PI			
	TAKEO MORIYA,TAKASHI ITO,YASUSHI SHINTANI,NOBUYUKI MIYAJIMA PC			
	C12N15/12,C12N1/21,C07K14/705,C07K16/28,C12P21/02,C12D1/68, PC			
	A61K38/00,			
	PC A61K45/00,A61K48/00,A61P1/00,A61P3/00,A61P9/00,A61P25/28, PC			
	A61P29/00,			
	PC A61P35/00,A61P37/00,G01N33/15,G01N33/50,G01N33/53,G01N33/566//			
	PC (C12N1/21,C12R1:19),(C12P21/02,C12R1:19)			
	CC Novel G protein-coupled receptor protein and its DNA. FH Key			
	Location/Qualifiers			
FT	source			
FT	1..1545			
FEATURES	Location/Qualifiers			
source	1..1545			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
BASE COUNT	429 a 320 c 290 g 506 t			
ORIGIN				
Query Match	68.1%;	Score 1543.4;	DB 6;	Length 1545;
Beet Local Similarity	99.9%;	Pred No. 0;		

[illegible]

1798 CTGGCTTTCTCATCATGTTGTTTCTATATTAATCTATGTTCTGTTCCATTCAAAAACC 1857
1081 CTGGCTTTCTCATCATGTTGTTTCTATATTAATCTATGTTCTGTTCCATTCAAAAACC 1140
1558 GCGTTGACAGCAGAGAGTAAGATGTTTGGAGAGAGTGGCTGTTGCAATCGT 1917
1141 GCGTTGACAGCAGAGAGTAAGATGTTTGGAGAGAGTGGCTGTTGCAATCGT 1200
1918 TTCTTTTATAGTGTCTGATGSCATCTGCGGATTCCTGTAATTGTTGTAATAATC 1977
1201 TTCTTTTATAGTGTCTGATGSCATCTGCGGATTCCTGTAATTGTTGTAATAATC 1260
1978 CTTTCCCTCTCCGGGTGGAATAACAGACACAATGATCTCTGATAGTATTTTTC 2037
1261 CTTTCCCTCTCCGGGTGGAATAACAGACACAATGATCTCTGATAGTATTTTTC 1320
2038 CTTCCAGTTAAACAGTCTTTGATTCATTCCTCTATCTCTCAACAACCACTTTTAAAG 2097
1321 CTTCCAGTTAAACAGTCTTTGATTCATTCCTCTATCTCTCAACAACCACTTTTAAAG 1380
2098 GACAGTTGAACAGCTGCTGACAAAGATGAGAGAGTAATGTAATTTTAAATAA 2157
1381 GACAGTTGAACAGCTGCTGACAAAGATGAGAGAGTAATGTAATTTTAAATAA 1440
2158 AAAAGTTTATCTACATTCATTTGTGTGATGATGAGAGCTCTCTCCCTGAACTTGGGGTT 2217
1441 AAAAGTTTATCTACATTCATTTGTGTGATGATGAGAGCTCTCTCCCTGAACTTGGGGTT 1500
Qy 2218 TTGAACAAATTAACACTTGGAGACAGTATTAATGAACCACTTTCC 2262
Db 1501 TTGAACAAATTAACACTTGGAGACAGTATTAATGAACCACTTTCC 1545

RESULT 14
BD141176 1473 bp DNA linear PAT 18-SEP-2002
LOCUS BD141176 1473 bp DNA linear PAT 18-SEP-2002
DEFINITION Novel G protein-coupled receptor protein and its DNA.
ACCESSION BD141176.1 GI:23236121
VERSION WO 0204640-A/5.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1473)
AUTHORS Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
TITLE Novel G protein-coupled receptor protein and its DNA
JOURNAL Takeda Chemical Industries Ltd, Takeo Moriya, Takashi Ito, Yasushi
Shintani, Nobuyuki Miyajima
OS Homo sapiens (human)
PN WO 0204640-A/5
PD 17-JUN-2002
PF 06-JUL-2001 WO 2001JP005878
PR 07-JUL-2000 JP 00P 211989, 18-DEC-2000 JP 00P 383794 PI
TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA PC
C12N15/12, C12N1/21, C07K14/705, C07K16/28, C12P21/02, C12Q1/68, PC
A61K38/00
PC A61K45/00, A61K48/00, A61P1/00, A61P3/00, A61P9/00, A61P25/28, PC
A61P29/00
PC A61P35/00, A61P37/00, G01N33/15, G01N33/50, G01N33/53, G01N33/566//
PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19)
CC Novel G protein-coupled receptor protein and its DNA. FH Key
Location/Qualifiers
FT source 1. 1473
Location/Qualifiers
1. 1473
Location/Qualifiers
/organism="Homo sapiens (human)".
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 405 a 310 c 278 g 480 t

ORIGIN
Query Match 61.3%; Score 1389.4; DB 6; Length 1473;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1472; Conservative 0; Mismatches 1; Indels 72; Gaps 1;
Qy 718 ATGGTATATACTACTTGAAGAGCTCTCCAGAGAGATGTGTGCCAAATGCCCAATC 777
Db 1 ATGGTATATACTACTTGAAGAGCTCTCCAGAGAGATGTGTGCCAAATGCCCAATC 60
Qy 778 AACTGGGTGATTTGGAGAGCATAGATTAATATCTCAAAATTTAGCTTCTGTG 837
Db 61 AACTGGGTGATTTGGAGAGCATAGATTAATATCTCAAAATTTAGCTTCTGTG 120
Qy 838 TGGGATTCGCTCAGAGTGTGTTCTGCTAGAAATCAATTTGGTTTGTCCAGAGAG 897
Db 121 TGGGATTCGCTCAGAGTGTGTTCTGCTAGAAATCAATTTGGTTTGTCCAGAGAG 138
Qy 898 ACATTTTCTTCATTAATAAATTTAGAGAACTGGATCTGTAGCAATAGATACGAG 957
Db 139 ACATTTTCTTCATTAATAAATTTAGAGAACTGGATCTGTAGCAATAGATACGAG 168
Qy 958 CTATCAGCTCACTTTTAAAGACTTGAAGCTTTTCAAAAAGCTGAACCTGTATCCAT 1017
Db 169 CTATCAGCTCACTTTTAAAGACTTGAAGCTTTTCAAAAAGCTGAACCTGTATCCAT 228
Qy 1018 CCGTTATGTATCTTCACAAGAACAGTTTGAAGTCTTAAACAACCTTCAAGTCTTAC 1077
Db 229 CCGTTATGTATCTTCACAAGAACAGTTTGAAGTCTTAAACAACCTTCAAGTCTTAC 288
Qy 1078 CTGAAAGATAGAGATTCCAATATTAACAACAGAAATGTTTCAACCAATGAAGATCTT 1137
Db 289 CTGAAAGATAGAGATTCCAATATTAACAACAGAAATGTTTCAACCAATGAAGATCTT 348
Qy 1138 TCTCACTTATTTCAAAAACCTTTCGATCTGCTCTATAGTCTCCCATGTCGAATATGT 1197
Db 349 TCTCACTTATTTCAAAAACCTTTCGATCTGCTCTATAGTCTCCCATGTCGAATATGT 408
Qy 1198 ATGCCCTTGAAGAGAGGAGATTTCTTCAATTTGAGAGAGCTTGGCTAAATATCTCTGA 1257
Db 409 ATGCCCTTGAAGAGAGGAGATTTCTTCAATTTGAGAGAGCTTGGCTAAATATCTCTGA 468
Qy 1258 ATATTTGTCTGGGTTATAGCTTTTCATTAACCTGCTTTGAAATCTTTTGTCTATTTGG 1317
Db 469 ATATTTGTCTGGGTTATAGCTTTTCATTAACCTGCTTTGAAATCTTTTGTCTATTTGG 528
Qy 1318 AGATCTTTCATTAAGCTGAATAATACACTCAAGCTATATGTCATCAAAATCTTTTGTGT 1377
Db 529 AGATCTTTCATTAAGCTGAATAATACACTCAAGCTATATGTCATCAAAATCTTTTGTGT 588
Qy 1378 GCTGATTCGCTGAGAGGATTTTACTTGTCTTTGTGGCAATTTGATTAATAATACCGA 1437
Db 589 GCTGATTCGCTGAGAGGATTTTACTTGTCTTTGTGGCAATTTGATTAATAATACCGA 648
Qy 1438 GGGGAGTATCAGAGATATGCTTGTCTGTGATGAGAGAGGCTGAGTCCGCTCATGAGG 1497
Db 649 GGGGAGTATCAGAGATATGCTTGTCTGTGATGAGAGAGGCTGAGTCCGCTCATGAGG 708
Qy 1498 TTCTGAGCATGCTGTCCACAGAGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1557
Db 709 TTCTGAGCATGCTGTCCACAGAGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 768
Qy 1558 AAGTCTGAGCATTTGTCTTCCCTTCAATTAACCTGAGTGAAGAGAGAGAGAGAGAG 1617
Db 769 AAGTCTGAGCATTTGTCTTCCCTTCAATTAACCTGAGTGAAGAGAGAGAGAGAGAG 828
Qy 1618 GTCACTCTCATTTGATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677
Db 829 GTCACTCTCATTTGATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
Qy 1678 AAGGATTAATTTGAGAACTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737
Db 889 AAGGATTAATTTGAGAACTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948


```

QY 1738 CAACAGAGATATTGGAAGCAAGGGATATTCTTGGAAATTTCTAGAGTGAACCTTG 1797
Db 949 CAAACAGAAAGATATTGGAAGCAAGGGATATTCTTGGAAATTTCTAGAGTGAACCTTG 1008
QY 1798 CTGGCTTTCTCATCTATTGGTTTCTCTATATAGTGTCTGCTCAATCAAAAAC 1857
Db 1009 CTGGCTTTCTCATCTATTGGTTTCTCTATATAGTGTCTGCTCAATCAAAAAC 1068
QY 1858 GCGCTTCAGACCAAGAGTAAGAAATTTGTTTGAAGAGAGGTGCTGTGCAAAATCGT 1917
Db 1069 GCGCTTCAGACCAAGAGTAAGAAATTTGTTTGAAGAGAGGTGCTGTGCAAAATCGT 1128
QY 1918 TTTCTTTTATATAGTGTCTCTGATGCGATCTGCTGATTCCTGTATTTTGAAGTAAATC 1977
Db 1129 TTTCTTTTATATAGTGTCTCTGATGCGATCTGCTGATTCCTGTATTTTGAAGTAAATC 1188
QY 1978 CTTTCCCTCTTCCGGGGGGAATATACAGACACATGACTCTCTGATGATGATTTTTC 2037
Db 1189 CTTTCCCTCTTCCGGGGGGAATATACAGACACATGACTCTCTGATGATGATTTTTC 1248
QY 2038 CTTTCCAGTTAAGAGTGTGTTGAATCCATCTCTATATCTCTCAAAACCACTTTTAAAG 2097
Db 1249 CTTTCCAGTTAAGAGTGTGTTGAATCCATCTCTATATCTCTCAAAACCACTTTTAAAG 1308
QY 2098 GACAAAGTTGAACAGCTGTCTGACCAAAACATCAGAGGAATCAATTTTCAAAATTTAAAAA 2157
Db 1309 GACAAAGTTGAACAGCTGTCTGACCAAAACATCAGAGGAATCAATTTTCAAAATTTAAAAA 1368
QY 2158 AAAAGTTTATCTATCATCTATGTTGTGATAGAGGACTCTCTCCCTGAAACTTGGGGTT 2217
Db 1369 AAAAGTTTATCTATCATCTATGTTGTGATAGAGGACTCTCTCCCTGAAACTTGGGGTT 1428
QY 2218 TTGAACAAATTAACACTTGGAGACAGTATATGAAACAGATTTC 2262
Db 1429 TTGAACAAATTAACACTTGGAGACAGTATATGAAACAGATTTC 1473

```

```

RESULT 15
BD141172 1830 bp DNA linear PAT 18-SEP-2002
LOCUS BD141172
DEFINITION Novel G protein-coupled receptor protein and its DNA.
ACCESSION BD141172
VERSION BD141172.1 GI:23236117
KEYWORDS NO 0204640-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Molecular: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1830)
Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
Novel G protein-coupled receptor protein and its DNA
Patent: WO 0204640-A 1 17-JAN-2002;
TAKEDA CHEMICAL INDUSTRIES LTD, TAKEO MORIYA, TAKASHI ITO, YASUSHI
SHINTANI, NOBUYUKI MIYAJIMA
OS Homo sapiens (human)
PN WO 0204640-A/1
PD 17-JAN-2002
PF 06-JUL-2001 WO 2001JP005878
PR 07-JUL-2000 JP ODP 211989,18-DEC-2000 JP ODP 383794 PI
TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA PC
C12N15/12, C12N1/21, C07K14/705, C07K16/28, C12P21/02, C12Q1/68, PC
A61K8/00,
PC A61K45/00, A61K48/00, A61P1/00, A61P3/00, A61P9/00, A61P25/28, PC
A61P23/00,
PC A61P35/00, A61P37/00, G01N33/15, G01N33/50, G01N33/53, G01N33/566//
PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19)
CC Novel G protein-coupled receptor protein and its DNA. FH Key

```

```

COMMENT
AUTHORS Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
TITLE Novel G protein-coupled receptor protein and its DNA
JOURNAL
FEATURES
FT source 1..1830 /organism='Homo sapiens (human)'.
Location/Qualifiers
1..1830
source

```

```

BASE COUNT 516 a 377 c 351 g 586 t
ORIGIN
Query Match 61.1%; Score 1384.8; DB 6; Length 1830;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 1828; Conservative 0; Mismatches 2; Indels 432; Gaps 1;
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
QY 1 ATGATGTTTCTGTTTAAACATCTCTTCAAGCTCAGATTGATTAAGATGTTCTT 60
Db 1 ATGATGTTTCTGTTTAAACATCTCTTCAAGCTCAGATTGATTAAGATGTTCTT 60
QY 61 CTACTTCATTTATGTTCTGATCAATGTCAAAGATTTTGCATCTGACCTCAAGTACATG 120
Db 61 CTACTTCATTTATGTTCTGATCAATGTCAAAGATTTTGCATCTGACCTCAAGTACATG 120
QY 121 ATCACTCTTCATGCAAAAAGATATTTTCCCTGAGGAATCTTACCAAGTCTTACC 180
Db 121 ATCACTCTTCATGCAAAAAGATATTTTCCCTGAGGAATCTTACCAAGTCTTACC 180
QY 181 CGAGCTTTTCACTGTGATGCAAGATGATCTGTGGAAAGGGGCGAGCAAGAGATCTGT 240
Db 181 CGAGCTTTTCACTGTGATGCAAGATGATCTGTGGAAAGGGGCGAGCAAGAGATCTGT 240
QY 241 GGTGACACTAGTGTGATGGGCGACATATTTTGGCAAGTGCATGGAATGCTTAAACGCTG 300
Db 241 GGTGACACTAGTGTGATGGGCGACATATTTTGGCAAGTGCATGGAATGCTTAAACGCTG 300
QY 301 GCTTATACAGAGAGGCTTTCTTAAACAGTATCCCAATGTGACTGCAAGAAACT 360
Db 301 GCTTATACAGAGAGGCTTTCTTAAACAGTATCCCAATGTGACTGCAAGAAACT 360
QY 361 GAATGGAATGTGTAATATGATGATTAAGTCTGTGCGCATGATTTCTTAAACATGTGACA 420
Db 361 GAATGGAATGTGTAATATGATGATTAAGTCTGTGCGCATGATTTCTTAAACATGTGACA 420
QY 421 TTAAGTCTCTTAAAGAAACCAAAATCCCAAGTCTTCCAGATTAAGTCTTCAATCAATC 480
Db 421 TTAAGTCTCTTAAAGAAACCAAAATCCCAAGTCTTCCAGATTAAGTCTTCAATCAATC 480
QY 481 ACAAACTTAAAGATATTTCTTCCAGATTAATGATTAAGACATATCCAGAAAGCA 540
Db 481 ACAAACTTAAAGATATTTCTTCCAGATTAATGATTAAGACATATCCAGAAAGCA 540
QY 541 TTTTGGATTAATGATCTGCAAAATTAATATTCACCAACCACTGATCACACCTTC 600
Db 541 TTTTGGATTAATGATCTGCAAAATTAATATTCACCAACCACTGATCACACCTTC 600
QY 497 ----- 496
Db 497 ----- 496
QY 601 AGACCTGGAATATTCAAAGACTTACATCAAGTAACTTGCTAATTTAGATGACAAATCCA 660
Db 601 AGACCTGGAATATTCAAAGACTTACATCAAGTAACTTGCTAATTTAGATGACAAATCCA 660
QY 497 ----- 496
Db 497 ----- 496
QY 661 ATACACGAATTCACAGGCTGTTTACGGGATTAATTCCTGTTTCTCTGTCTATG 720
Db 661 ATACACGAATTCACAGGCTGTTTACGGGATTAATTCCTGTTTCTCTGTCTATG 720
QY 497 ----- 496
Db 497 ----- 496
QY 721 GTTAATTAATTAATTAAGAGCTCTTCCCAAGCAGATGTGCCCCAATGCTCACTCAAC 780
Db 721 GTTAATTAATTAATTAAGAGCTCTTCCCAAGCAGATGTGCCCCAATGCTCACTCAAC 780
QY 497 ----- 496
Db 497 ----- 496
QY 781 TGGGTGATTTGGAAGCAATAGATTAAGTATCTCAAAATTTCTAGTTCTGTGTGC 840
Db 781 TGGGTGATTTGGAAGCAATAGATTAAGTATCTCAAAATTTCTAGTTCTGTGTGC 840
QY 497 ----- 496
Db 497 ----- 496
QY 841 GATTCGCTACAGTGTGTTTCTGCTAGAAATCAATGTTTGTTCAGAGAAACA 900
Db 841 GATTCGCTACAGTGTGTTTCTGCTAGAAATCAATGTTTGTTCAGAGAAACA 900
QY 497 ----- 496
Db 497 ----- 496
QY 901 TTTTCTTCATTAATAAATTTAGAGAACTGATCTGTCTTACGAATACGATTAACGAGACTA 960
Db 901 TTTTCTTCATTAATAAATTTAGAGAACTGATCTGTCTTACGAATACGATTAACGAGACTA 960
QY 497 ----- 496
Db 497 ----- 496

```

Qy 961 TCACCTCACCTTTTAAAGCTTGAAGCTTCTACAAAAGCTGAACCTGTCAATCCCT 1020
Db 529 TCACCTCACCTTTTAAAGCTTGAAGCTTCTACAAAAGCTGAACCTGTCAATCCCT 588
Qy 1021 CTATATATCTTCAACAGAACCAAGTTGAAAGCTTTAAACAATCTAGTCTCTAGACCTG 1080
Db 589 CTATATATCTTCAACAGAACCAAGTTGAAAGCTTTAAACAATCTAGTCTCTAGACCTG 648
Qy 1081 GAAAGATAGAGATTCCAAATATATAACACAGAAATGTTTCAACCCATGAAATCTTTCT 1140
Db 649 GAAAGATAGAGATTCCAAATATATAACACAGAAATGTTTCAACCCATGAAATCTTTCT 708
Qy 1141 CACATTTATTTCAAAAACCTTGCATATCTGTCTATAGCTCCCAATGCTCGAATATATG 1200
Db 709 CACATTTATTTCAAAAACCTTGCATATCTGTCTATAGCTCCCAATGCTCGAATATATG 768
Qy 1201 CCCTTGACGAGCGGCAATTTCTTCAATTGAGACCTCTTGAGTAAACAATATCTGAGAA 1260
Db 769 CCCTTGACGAGCGGCAATTTCTTCAATTGAGACCTCTTGAGTAAACAATATCTGAGAA 828
Qy 1261 TTTGTCTGGGTTATAGCTTTTCAATACCTGCTTTGGAATCTTTTGTATGTCATGAGA 1320
Db 829 TTTGTCTGGGTTATAGCTTTTCAATACCTGCTTTGGAATCTTTTGTATGTCATGAGA 888
Qy 1321 TCTTTCATTTAAGCTGAATAATACAACCTACGCTATGCTCATCAAAAATCTTTGTGTGCT 1380
Db 889 TCTTTCATTTAAGCTGAATAATACAACCTACGCTATGCTCATCAAAAATCTTTGTGTGCT 948
Qy 1381 GATTGCTGATGGGTGTTTACTTGTCTTTGTCATTTTGAGATATATAATACCGAGGG 1440
Db 949 GATTGCTGATGGGTGTTTACTTGTCTTTGTCATTTTGAGATATATAATACCGAGGG 1008
Qy 1441 CAGATACGAAGTATGCTGTGTGAGATGAGACGCGGCTGATGGGGTTC 1500
Db 1009 CAGATACGAAGTATGCTGTGTGAGATGAGACGCGGCTGATGGGGTTC 1068
Qy 1501 CTGGCAGTCTGTCCACCGAAGTCTCTGTCTGTACTGACTTGAATCTTTGAGAGAAG 1560
Db 1069 CTGGCAGTCTGTCCACCGAAGTCTCTGTCTGTACTGACTTGAATCTTTGAGAGAAG 1128
Qy 1561 TTCTGTGATATGTCTTCCCTTCAATACCTGACCTGGAAGCGGACACTCACTC 1620
Db 1129 TTCTGTGATATGTCTTCCCTTCAATACCTGACCTGGAAGCGGACACTCACTC 1188
Qy 1621 ATCCATTTGATCTGATGCGCGGATTTTAAATAGCTATACTCAATTTGGAATAG 1680
Db 1189 ATCCATTTGATCTGATGCGCGGATTTTAAATAGCTATACTCAATTTGGAATAG 1248
Qy 1681 GATTATTTTGAACCTTTATGGAATAATGAGATGTTTCCACTTATATGACCA 1740
Db 1249 GATTATTTTGAACCTTTATGGAATAATGAGATGTTTCCACTTATATGACCA 1308
Qy 1741 ACAGAAATATTTGAAGCAAGGATTTCTCTTGAATTTCTTGAATTTCTTGAATTTCTGCTG 1800
Db 1309 ACAGAAATATTTGAAGCAAGGATTTCTCTTGAATTTCTTGAATTTCTTGAATTTCTGCTG 1368
Qy 1801 GCTTTCTCATATGTTTCTTATATTAATATGTTCTGTTCAATCAAAAACCGCC 1860
Db 1369 GCTTTCTCATATGTTTCTTATATTAATATGTTCTGTTCAATCAAAAACCGCC 1428
Qy 1861 TTGACAGCACAGAGATAGAAATGTTTGAAGAGAGTGGCTGTGCAATCGTTTC 1920
Db 1429 TTGACAGCACAGAGATAGAAATGTTTGAAGAGAGTGGCTGTGCAATCGTTTC 1488
Qy 1921 TTTTATATAGTGTCTCTGATGCTGCTGATTTCTGTATTTGTATGTTAAATCTTT 1980
Db 1489 TTTTATATAGTGTCTCTGATGCTGCTGATTTCTGTATTTGTATGTTAAATCTTT 1548
Qy 1981 TCCCTCTTCCGGGTGGAATAACAGACAAATGACTTCTGATATGTTTCTTTCTT 2040
Db 1549 TCCCTCTTCCGGGTGGAATAACAGACAAATGACTTCTGATATGTTTCTTTCTT 1608

Qy 2041 CCAATTACAGTCTTGAATCCATCTCTATATCTCAACAACAACTTTTAAAGAC 2100
Db 1609 CCAATTACAGTCTTGAATCCATCTCTATATCTCAACAACAACTTTTAAAGAC 1668
Qy 2101 AAGTTGAAACAGCTGCTGCAAAACATCAGAGGAAATCAATTTTCAAAAATTTAAAAA 2160
Db 1669 AAGTTGAAACAGCTGCTGCAAAACATCAGAGGAAATCAATTTTCAAAAATTTAAAAA 1728
Qy 2161 AGTTTATCTATCCATTGATGATGAGAACTGCTCTTCCCTGAAAATTTGGGTTTGG 2220
Db 1729 AGTTTATCTATCCATTGATGATGAGAACTGCTCTTCCCTGAAAATTTGGGTTTGG 1788
Qy 2221 AACAAAATTAACATTTGAGACAGATATATGAAAACAGTTTC 2262
Db 1789 AACAAAATTAACATTTGAGACAGATATATGAAAACAGTTTC 1830

Search completed: November 28, 2003, 18:06:00
Job time : 8055 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2003, 15:41:35 ; Search time 582 Seconds
(without alignments)
10505.545 Million cell updates/sec

Title: US-09-928-175-1

Perfect score: 2265

Sequence: 1 atgattgtttcttcggtttc.....taatgaacacagttctctag 2265

Scoring table: IDENTITY NUC
Gapox 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2265	100.0	2339	25	AA151335
2	2258.8	99.7	2262	24	ABL40196
3	2214	97.7	2214	24	ABK51944
4	2104.8	92.9	2190	24	ABL40197
5	2060	90.9	2142	24	ABK51947
6	2002.2	88.4	2181	24	AAFP8580
7	1543.4	68.1	1545	24	ABL40191
8	1389.4	61.3	1473	24	ABL40192

9	1384.8	61.1	1830	24	ABL40188	Human G protein-co
10	1105.8	48.8	1121	25	ABT33349	NOVX DNA sequence
11	1065.4	47.1	1068	22	AA507943	Human G protein-co
12	1065	47.0	1065	24	ABL40194	Human G protein-co
13	735	32.5	2274	25	AB242570	Human G protein-co
14	735	32.5	2467	20	AA225345	Human LGR7 long fo
15	686.2	30.3	3584	24	AA225346	Human LGR7 short f
16	547	24.2	575	24	AA205498	Human FSH-like GPC
17	496.8	21.9	1804	21	AA290524	Human G-protein
18	496.8	21.9	1804	24	ABX73054	Human G-protein co
19	482.2	21.3	1191	22	AA06507	Human G-protein co
20	411	18.1	429	24	AA598139	Human DNA for pote
21	370	16.3	592	22	AA546861	Human G protein-co
22	370	16.3	592	24	ABK61639	CDNA encoding nove
23	321	14.2	321	24	AA032027	Human novel G-prot
24	266.4	11.8	760	22	AA135687	Human musculoskele
25	266.4	11.8	760	25	ABX58675	CDNA encoding nove
26	232.4	10.3	576	22	AA546918	Human G protein-co
27	232.4	10.3	576	24	ABK61696	CDNA encoding nove
28	227.6	10.0	1126	25	ACA04807	CDNA encoding huma
29	226.4	10.0	572	24	AB190585	Human polynucleoti
30	222	9.8	612	24	ABX73081	Human GPCR CDNA #4
31	221	9.8	1018	22	AAH51001	Human nGPCR57 codi
32	221	9.8	1018	24	ABX70234	DNA encoding human
33	185	8.2	562	24	ABX73082	Human GPCR CDNA #4
34	184.4	8.1	530	22	AA195984	Human expressed po
35	184.4	8.1	530	22	AA135644	Human musculoskele
36	184.4	8.1	530	22	ABA06471	Human CDNA SEQ ID
37	184.4	8.1	530	22	AA528950	CDNA encoding for
38	184.4	8.1	530	22	AA529573	Human endocrine po
39	184.4	8.1	530	22	AA530187	DNA encoding rena
40	184.4	8.1	530	22	AA530845	CDNA encoding nove
41	184.4	8.1	530	23	ABK43875	DNA encoding nove
42	184.4	8.1	530	24	ABK83808	Human polynucleoti
43	184.4	8.1	530	24	ABT07785	Novel human nuclei
44	184.4	8.1	530	25	ABX58632	CDNA encoding nove
45	183.2	8.1	1089	22	AA199557	Human expressed po

ALIGNMENTS

RESULT 1	AA151335	standard; DNA; 2339 BP.
ID	AA151335	
AC	AA151335;	
XX		
DT	20-MAR-2003	(first entry)
XX		
DE	Human LGR9 G-protein coupled receptor (GPCR) coding sequence.	
XX		
KW	Human; gene; ds; gene therapy; LGR9; G-protein coupled receptor; GPCR;	
KW	muscle mass decrease; muscle atrophy; cryptorchidism; inflammation;	
KW	asthma; allergy; oedema; high blood pressure; low blood pressure.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	71..2335
FT		/tag= a
FT		/product= "Human LGR9 G-protein coupled receptor (GPCR)"
XX		
PN	MO2002102992-A2.	
XX		
PD	27-DEC-2002.	
XX		
PF	19-JUN-2002; 2002WO-US19528.	
XX		
PR	19-JUN-2001; 2001US-299385P.	
XX		
PA	(REGC-) REGENERON PHARM INC.	
XX		

PI Murphy AJ, Glass DJ;
XX WPI; 2003-157041/15.
DR P-PSDB; AAO16303.
XX
PT New human LGR9 G-protein coupled receptor (GPCR) polypeptide, useful
for treating or diagnosing disorders associated with altered LGR9 GPCR
expression, e.g. muscle atrophy, cryptorchidism or inflammation, and in
pharmacogenomics
XX
PS Disclosure; Fig 1A-F; 131dp; English.
XX
CC The invention comprises the amino acid and coding sequence of the human
LGR9 G-protein coupled receptor (GPCR). The human LGR9 GPCR is useful in
diagnosing and treating conditions associated with a decrease in muscle
mass or atrophy. The human LGR9 GPCR DNA and protein sequences of the
invention are useful for treating diseases associated with altered human
LGR9 GPCR expression, such as: cryptorchidism; asthma; inflammation;
allergy; oedema; and high or low blood pressure. The present DNA sequence
encodes the human LGR9 GPCR of the invention.
CC
XX
SQ Sequence 2339 BP; 676 A; 482 C; 435 G; 746 T; 0 other;
Query Match 100.0%; Score 2265; DB 25; Length 2339;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGATGTTTTCGTTTAAACATCTCTTCAGCTCAGATTGATTACATGTTCTT 60
DB 71 ATGATGTTTTCGTTTAAACATCTCTTCAGCTCAGATTGATTACATGTTCTT 130
QY 61 CTACTTCATTTCATCGTTCTGATCATATGTCAGAGATTTTGACACTGACCTCAAGTAGCATG 120
DB 131 CTACTTCATTTCATCGTTCTGATCATATGTCAGAGATTTTGACACTGACCTCAAGTAGCATG 190
QY 121 ATCACTCTTCAGTCCAAAAGATATTTTCCGTGGGAATCTTACCAAGTCTTACC 180
DB 191 ATCACTCTTCAGTCCAAAAGATATTTTCCGTGGGAATCTTACCAAGTCTTACC 250
QY 181 CGAGCTTTTCACTGTGATGGCAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 240
DB 251 CGAGCTTTTCACTGTGATGGCAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 310
QY 241 GGTACACTGATGATGGGCGACATATTTGGCAAGTGCATGGAATGCTAACAGGCTG 300
DB 311 GGTACACTGATGATGGGCGACATATTTGGCAAGTGCATGGAATGCTAACAGGCTG 370
QY 301 GCCTTAAACAGAGAGTCTTCTTAAACAGATCCACAAATGCTGTGACTGCAAGAACT 360
DB 371 GCCTTAAACAGAGAGTCTTCTTAAACAGATCCACAAATGCTGTGACTGCAAGAACT 430
QY 431 GAAATGGGAATGTGTAATGTGACTTAAAGTCTGTGCGAGATGATTTCTAACATGTGACA 420
DB 491 TTAATGGGAATGTGTAATGTGACTTAAAGTCTGTGCGAGATGATTTCTAACATGTGACA 490
QY 421 TTAATGGTCTCTTAAAGAAACAAATCCAGAGTCTTCCAGATTAAGTTTATGAAATAC 480
DB 491 TTAATGGTCTCTTAAAGAAACAAATCCAGAGTCTTCCAGATTAAGTTTATGAAATAC 550
QY 481 ACAAAACCTTAAAGATATTTCTTCAAGATATGACTTAAAGATATGACATATCCAGAGAACT 540
DB 551 ACAAAACCTTAAAGATATTTCTTCAAGATATGACTTAAAGATATGACATATCCAGAGAACT 610
QY 541 TTTTGGATTAATGTAATGTGCAAAATTAATATCTCAACCAACCTGATCAACCTCTC 600
DB 611 TTTTGGATTAATGTAATGTGCAAAATTAATATCTCAACCAACCTGATCAACCTCTC 670
QY 601 AGACCTGGAATATCAAGAGCTTACATGAGCTTAAGCTTAACTTCTTGAATGACAACTCA 660
DB 671 AGACCTGGAATATCAAGAGCTTACATGAGCTTAACTTCTTGAATGACAACTCA 730
QY 661 ATAAACAGAAATTTCAAGAGCTTGTGTTACGGGATTAATTCCTGTTTCTGTCTATG 720

DB 731 ATAAACAGAAATTTCAAGAGCTTGTGTTACGGGATTAATTCCTGTTTCTGTCTATG 790
QY 721 GTTAAATTAATCTTAAGAGCTCTTCCCAAGAGATGTGTGCCAAATGCTCACTCAAC 780
DB 791 GTTAAATTAATCTTAAGAGCTCTTCCCAAGAGATGTGTGCCAAATGCTCACTCAAC 850
QY 781 TGGGAGATTTTGAAGGCAATGAAATTAAGTATCTCAAAATTTCTACGTTTCTGTCTG 840
DB 851 TGGGAGATTTTGAAGGCAATGAAATTAAGTATCTCAAAATTTCTACGTTTCTGTCTG 910
QY 841 GATTCGCTCAGAGTGTGTTCTGCTAGAAATCAATGTTGTTGTCAGAGAGACA 900
DB 911 GATTCGCTCAGAGTGTGTTCTGCTAGAAATCAATGTTGTTGTCAGAGAGACA 970
QY 901 TTTTCTTCATTAATAAATTTAGAGAACTGATCTGTCTAGCAATACGATTAACGAGCTA 960
DB 971 TTTTCTTCATTAATAAATTTAGAGAACTGATCTGTCTAGCAATACGATTAACGAGCTA 1030
QY 961 TCACCTCAGCTTTTAAACATTTGAAGCTTCTACAAAACCTGAACCTGTCAATCCCT 1020
DB 1031 TCACCTCAGCTTTTAAACATTTGAAGCTTCTACAAAACCTGAACCTGTCAATCCCT 1090
QY 1021 CTTATGATCTTCAACAAGAACAGTTGAAAGCTTAAACAACTTCACTCTAGACCTG 1080
DB 1091 CTTATGATCTTCAACAAGAACAGTTGAAAGCTTAAACAACTTCACTCTAGACCTG 1150
QY 1081 GAAAGATAGAGATTCCAATATTAACACAGATTTTCAACCCATGAGAAATCTTTCT 1140
DB 1151 GAAAGATAGAGATTCCAATATTAACACAGATTTTCAACCCATGAGAAATCTTTCT 1210
QY 1141 CACATTTATTTCAAAAACCTTGATCTGCTCTGATGCTCCCAATGCGAATATGATG 1200
DB 1211 CACATTTATTTCAAAAACCTTGATCTGCTCTGATGCTCCCAATGCGAATATGATG 1270
QY 1201 CCCTTACGAGACGGCATTTCTTCAATTTGAGAACCTCTGCTTAACAATATCTCAGAAAT 1260
DB 1271 CCCTTACGAGACGGCATTTCTTCAATTTGAGAACCTCTGCTTAACAATATCTCAGAAAT 1330
QY 1261 TTTGCTGGGTTATGACCTTCAATACCTGTTTGGAAATCTTTTGTGATGGCAATGAA 1320
DB 1331 TTTGCTGGGTTATGACCTTCAATACCTGTTTGGAAATCTTTTGTGATGGCAATGAA 1390
QY 1321 TCTTTCAATTAACCTGAAATACAACTACAGCTATGCTCAAAATCTTTGTTGCT 1380
DB 1391 TCTTTCAATTAACCTGAAATACAACTACAGCTATGCTCAAAATCTTTGTTGCT 1450
QY 1381 GATTCGCTGATGGGTTTACTGTTCTTGTGGAATTTTGGATTAATAATCCGAGGG 1440
DB 1451 GATTCGCTGATGGGTTTACTGTTCTTGTGGAATTTTGGATTAATAATCCGAGGG 1510
QY 1441 CAGATTCAGAAATGATGCTTGTGCTGTGATGGAAGGCTGAGTCCGCTCATGGGGTTC 1500
DB 1511 CAGATTCAGAAATGATGCTTGTGCTGTGATGGAAGGCTGAGTCCGCTCATGGGGTTC 1570
QY 1501 CTGGGCAATGCTGTCACACGAACTCTCTGTTCTGCTACTGACTACTTGAAGGAG 1560
DB 1571 CTGGGCAATGCTGTCACACGAACTCTCTGTTCTGCTACTGACTACTTGAAGGAG 1630
QY 1561 TTCCTGCTCAATGCTTTCCTTCACTGATCACTTGAACCTGGAACCGGCACTCATGTC 1620
DB 1631 TTCCTGCTCAATGCTTTCCTTCACTGATCACTTGAACCTGGAACCGGCACTCATGTC 1690
QY 1621 ATCTCTATTTGCAATCTGGAATGGGGGATTTTAAATAGCTTAATTCATTTTGAATTAAG 1680
DB 1691 ATCTCTATTTGCAATCTGGAATGGGGGATTTTAAATAGCTTAATTCATTTTGAATTAAG 1750
QY 1681 GATTAATTTGAAACCTTTATGGAATAATGAGATATGTTTCCACTTAATTAATGACAA 1740
DB 1751 GATTAATTTGAAACCTTTATGGAATAATGAGATATGTTTCCACTTAATTAATGACAA 1810
QY 1741 ACAGAGATATTTGAGACCAAGGGTATCTCTTGAATTTTCTGAGTGTGAATCTGCTG 1800
DB 1811 ACAGAGATATTTGAGACCAAGGGTATCTCTTGAATTTTCTGAGTGTGAATCTGCTG 1870

Oy	1801	GCTTTTCACATCATGTGTTTTCCATATATCTATGCTCTGTCCATTCAAAAACGGCC	1866
Db	1871	GCTTTTCACATCATGTGTTTTCCATATATCTATGCTCTGTCCATTCAAAAACGGCC	1930
Oy	1861	TTGCAGACCAAGAAAGTAAGAAATTGTTTTGGAGAGAGGTGGCTGTTGCCAAATCGTTTC	1920
Db	1931	TTGCAGACCAAGAAAGTAAGAAATTGTTTTGGAGAGAGGTGGCTGTTGCCAAATCGTTTC	1990
Oy	1921	TTTTTTATAGTGTTCTCGATGCGCATCTGCTGGATTCCTGTAATTTGTATGTTAAATCCTT	1980
Db	1991	TTTTTTATAGTGTTCTCGATGCGCATCTGCTGGATTCCTGTAATTTGTATGTTAAATCCTT	2050
Oy	1981	TCCTCTTCGGGGTGGAAATACCAAGACAACAATGACTTCCTGGATATGTAATTTTTTCCTT	2040
Db	2051	TCCTCTTCGGGGTGGAAATACCAAGACAACAATGACTTCCTGGATATGTAATTTTTTCCTT	2110
Oy	2041	CCAGTTACACGTGCTTGAATCCAAATCCTTAATCTTCACAAACAATTTTTTAAAGAC	2100
Db	2111	CCAGTTACACGTGCTTGAATCCAAATCCTTAATCTTCACAAACAATTTTTTAAAGAC	2170
Oy	2101	AAGTTGAACAACGTGCTGCACAAACATCAGAGGAATCAATTTTTCAAAATTTAAAAAAA	2160
Db	2171	AAGTTGAACAACGTGCTGCACAAACATCAGAGGAATCAATTTTTCAAAATTTAAAAAAA	2230
Oy	2161	AGTTTATCTACATCCATTGTGTGATTAAGAGACTCTCTTCCCTGGAACCTTGGGGTTTTG	2220
Db	2231	AGTTTATCTACATCCATTGTGTGATTAAGAGACTCTCTTCCCTGGAACCTTGGGGTTTTG	2290
Oy	2221	AACAAAATPACACTTGAGACAGATTAATGAACAAGTTTCTTAG	2265
Db	2291	AACAAAATPACACTTGAGACAGATTAATGAACAAGTTTCTTAG	2335

Result	ID	ABL40196	standard; cDNA; 2262 BP.
AC	ABL40196;		
DT	23-MAY-2002	(first entry)	
DE	Human G protein-coupled receptor TGR17-5 encoding cDNA SEQ ID NO:14.		
KW	Human; G protein-coupled receptor; TGR17-5; neotropic; antiinflammatory; vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy; neurological; inflammatory; circulatory; degenerative; immune system; digestive disease; cancer; gene; ss.		
OS	Homo sapiens.		
PH	Key	location/Qualifiers	
FT	CDS	1..2262	
FT		/*tag= a	
FT		/partial	
FT		/product= "TGR17-5"	
FT		/note= "no stop codon given"	
PN	WO200204640-A1.		
PD	17-JAN-2002.		
PF	06-JUL-2001; 2001WO-JP05878.		
PR	07-JUL-2000; 2000JP-0231989.		
PR	18-DEC-2000; 2000JP-0383794.		
PA	(TAKE) TAKEDA CHEM IND LTD.		
PI	Moriya T, Ito T, Shintani Y, Miyajima N;		
DR	WPI, 2002-179706/23.		
DR	P-PSDB; ABB06254.		

xx	G-protein coupled receptor protein TGR17 of human origin and DNA
pt	encoding it for diagnosis and treatment of cancer and circulatory and
xx	other diseases associated with its expression -
xx	Claim 7; Page 130-131; 145pp; Japanese.
ps	
xx	The present invention describes a human guanine nucleotide binding
cc	protein (G protein) coupled receptor protein designated TGR17, which has
cc	5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
cc	additional residues at the N-terminal. The TGR17 proteins have nootropic,
cc	antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
cc	The TGR17 polynucleotides and protein can be used in gene therapy and
cc	protein therapy. G-protein coupled receptor proteins are cell membrane
cc	proteins mediating the cellular response to a large variety of signalling
cc	molecules. The TGR17 polynucleotides and proteins can be used in the
cc	diagnosis, treatment and prevention of diseases including neurological,
cc	inflammatory, circulatory, degenerative, immune system and digestive
cc	diseases and cancer. The present sequence encodes human TGR17-5 from the
cc	present invention.
cc	
sq	Sequence 2262 BP; 650 A; 467 C; 421 G; 724 T; 0 other;
	Query Match 99.7%; Score 2258.8; DB 24; Length 2262;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 2260; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
Qy	1 ATGATGTGTTTTCGTTTAAACATCTCTACAGCTCAGATTGATTACATGTTCTT 60
Db	1 ATGATGTGTTTTCGTTTAAACATCTCTACAGCTCAGATTGATTACATGTTCTT 60
Qy	61 CTACTTCATTCATCGTTCGTATCATCATGTCAAAGATTTTGGACTGACTCAAGTAGCATG 120
Db	61 CTACTTCATTCATCGTTCGTATCATCATGTCAAAGATTTTGGACTGACTCAAGTAGCATG 120
Qy	121 ATCACTCCTTCATGCGCAAAAGATATTTTCCCTGTGGGAATCTTACCAAGCTTACCC 180
Db	121 ATCACTCCTTCATGCGCAAAAGATATTTTCCCTGTGGGAATCTTACCAAGCTTACCC 180
Qy	181 CGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 240
Db	181 CGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 240
Qy	241 GGTGACACTAGTGAATGGGCGACCATATTTGGCACATGTCATGGAATGCTAACGCGTG 300
Db	241 GGTGACACTAGTGAATGGGCGACCATATTTGGCACATGTCATGGAATGCTAACGCGTG 300
Qy	301 GCCTTAAACACGAGAGTGTCTTTCTTAAACAGTATCCACAAATGCTGACGCAAAAGAACT 360
Db	301 GCCTTAAACACGAGAGTGTCTTTCTTAAACAGTATCCACAAATGCTGACGCAAAAGAACT 360
Qy	361 GAATTGGAATGTGAATGATGACTTAAAGTCTGTGCCGATGATTTCTAACATGTGACA 420
Db	361 GAATTGGAATGTGAATGATGACTTAAAGTCTGTGCCGATGATTTCTAACATGTGACA 420
Qy	421 TTACTGTCTCTTAAGAAAAACAAATCCACAGTCTTCCAGATTAAGTTTTCATCAATTC 480
Db	421 TTACTGTCTCTTAAGAAAAACAAATCCACAGTCTTCCAGATTAAGTTTTCATCAATTC 480
Qy	481 ACAAAACTTAAAGATATTTCTGAGATATTTGATTAACATATCCAGAAAGACA 540
Db	481 ACAAAACTTAAAGATATTTCTGAGATATTTGATTAACATATCCAGAAAGACA 540
Qy	541 TTTTGTGATTAATGTAAATCTGCAATATTAATATCTCAACCAACATGCAATCAACCTTC 600
Db	541 TTTTGTGATTAATGTAAATCTGCAATATTAATATCTCAACCAACATGCAATCAACCTTC 600
Qy	601 AGACCTGGAATATTTCAAAAGCTTACATCATGCTTAATTTGGCTTAATCTAGATACATTC 660
Db	601 AGACCTGGAATATTTCAAAAGCTTACATCATGCTTAATTTGGCTTAATCTAGATACATTC 660
Qy	661 ATAAACGAATTTCAACAGCGCTGTGTTTACGGGATTTAAATTCCTGTGTTTCTGTCTATG 720
Db	661 ATAAACGAATTTCAACAGCGCTGTGTTTACGGGATTTAAATTCCTGTGTTTCTGTCTATG 720

```

Db 661 ATACACAGAAATTCACAGCGCTGTTAGGGAATTAATTCCTGTTTCCGTCTATG 720
Qy 721 GTTAATTAAGCTCTTCCAGACAGATGTGCCCCAAATGCTCACTCAAC 780
Db 721 GTTAATTAAGCTCTTCCAGACAGATGTGCCCCAAATGCTCACTCAAC 780
Qy 781 TGGGTGATTTGGAGGCAATGAATTAAGTATCTCAAAATCTACGTTTCTGTGTCG 840
Db 781 TGGGTGATTTGGAGGCAATGAATTAAGTATCTCAAAATCTACGTTTCTGTGTCG 840
Qy 841 GATTCGCTCAGAGTGTGTTCTGCTAGAAATCAATGGTTTGTTCAGAGAGACA 900
Db 841 GATTCGCTCAGAGTGTGTTCTGCTAGAAATCAATGGTTTGTTCAGAGAGACA 900
Qy 901 TTTTCTTCATTAATAAATTTAGAGAACTGGATCTGTCTAGCAATAAGTAACGAGCTA 960
Db 901 TTTTCTTCATTAATAAATTTAGAGAACTGGATCTGTCTAGCAATAAGTAACGAGCTA 960
Qy 961 TCACCTCACCTTTTAAAGCTTGAAGCTTCTACAAAGCTGAACCTGTCAATCCAT 1020
Db 961 TCACCTCACCTTTTAAAGCTTGAAGCTTCTACAAAGCTGTCAATCCAT 1020
Qy 1021 CTATATATCTTCAACAAGAACAGTTGAAAGCTTAAACAATCTGAGTCTAGAGCTG 1080
Db 1021 CTATATATCTTCAACAAGAACAGTTGAAAGCTTAAACAATCTGAGTCTAGAGCTG 1080
Qy 1081 GAAAGATAGAGATTCCTCAATATATAACAAGATGTTTCAACCCATGAAGATCTTTCT 1140
Db 1081 GAAAGATAGAGATTCCTCAATATATAACAAGATGTTTCAACCCATGAAGATCTTTCT 1140
Qy 1141 CACATTTATTTCAAAAACCTTGATCTGCTCTATGCTCCCATGTCGGAATATGATG 1200
Db 1141 CACATTTATTTCAAAAACCTTGATCTGCTCTATGCTCCCATGTCGGAATATGATG 1200
Qy 1201 CCCTTGACGAGCGCAATTCCTCATTTGAGAGACCTTGCTTAACAATATCTCAGATA 1260
Db 1201 CCCTTGACGAGCGCAATTCCTCATTTGAGAGACCTTGCTTAACAATATCTCAGATA 1260
Qy 1261 TTTGTCTGGGTATTAAGCTTCTCATTAAGCTTGGAAATTTTGTGCTATGAGAGA 1320
Db 1261 TTTGTCTGGGTATTAAGCTTCTCATTAAGCTTGGAAATTTTGTGCTATGAGAGA 1320
Qy 1321 TCTTTCAATTAAGCTGAAATACACTGATGCTATGCTCAAAATCCTTTGTGCT 1380
Db 1321 TCTTTCAATTAAGCTGAAATACACTGATGCTATGCTCAAAATCCTTTGTGCT 1380
Qy 1381 GATTCCTGATGGGTGTTTACTGTTCTTTGTTGGCAATTTTCAATTAATAACGAGG 1440
Db 1381 GATTCCTGATGGGTGTTTACTGTTCTTTGTTGGCAATTTTCAATTAATAACGAGG 1440
Qy 1441 CAGTATCAAGAAATAGCTTGTGCTGATGAGAGAGCTGAGCGGCTCATGGGGTTC 1500
Db 1441 CAGTATCAAGAAATAGCTTGTGCTGATGAGAGAGCTGAGCGGCTCATGGGGTTC 1500
Qy 1501 CTGGCAATGCTGCTCAACCGAAGCTCTGTTCTGCTACTGACTACTTGTGAGAA 1560
Db 1501 CTGGCAATGCTGCTCAACCGAAGCTCTGTTCTGCTACTGACTACTTGTGAGAA 1560
Qy 1561 TTTCTGCTGATGTTCTTCCCTTCAATTAAGCTTGAATTAAGCTTGAATTAAG 1620
Db 1561 TTTCTGCTGATGTTCTTCCCTTCAATTAAGCTTGAATTAAGCTTGAATTAAG 1620
Qy 1621 ATCTCATTTGCAATCTGATGAGGAGGATTTTAATTAAGCTTGAATTAAG 1680
Db 1621 ATCTCATTTGCAATCTGATGAGGAGGATTTTAATTAAGCTTGAATTAAG 1680
Qy 1681 GATTAATTTGAAACCTTTTAAGGAAATGAGTATGTTTCCCACTTTATTAAGCA 1740
Db 1681 GATTAATTTGAAACCTTTTAAGGAAATGAGTATGTTTCCCACTTTATTAAGCA 1740
Qy 1741 ACAGAAATATTTGAGAGCAAGGATTTCTTGTGAATTTTCTGAGTGTGAACCTGCTG 1800
Db 1741 ACAGAAATATTTGAGAGCAAGGATTTCTTGTGAATTTTCTGAGTGTGAACCTGCTG 1800

```

```

Qy 1801 GCTTTCTCATGATGTTGTTTCTCTATATTAATTAATGTTCTGTTCAATCAAAAACCGCC 1860
Db 1801 GCTTTCTCATGATGTTGTTTCTCTATATTAATTAATGTTCTGTTCAATCAAAAACCGCC 1860
Qy 1861 TTGACAGACACAGAGTAAAGAAATGTTTGGAGAGAGTGGCTGTGCAAAATGCTTTC 1920
Db 1861 TTGACAGACACAGAGTAAAGAAATGTTTGGAGAGAGTGGCTGTGCAAAATGCTTTC 1920
Qy 1921 TTTTATTAAGTGTCTGATGTCATCTGCTGATTCCTGTAATTTGATTAATTCCTT 1980
Db 1921 TTTTATTAAGTGTCTGATGTCATCTGCTGATTCCTGTAATTTGATTAATTCCTT 1980
Qy 1981 TCCCTCTTCCGGGTGAAATATACAGACACATGACTTCTGATATGATTTTTCCTT 2040
Db 1981 TCCCTCTTCCGGGTGAAATATACAGACACATGACTTCTGATATGATTTTTCCTT 2040
Qy 2041 CCAGTTAACAGTCTTTGATTCATTCCTTACTCTCAACAACCACTTTTAAAGAC 2100
Db 2041 CCAGTTAACAGTCTTTGATTCATTCCTTACTCTCAACAACCACTTTTAAAGAC 2100
Qy 2101 AAGTTGAAACAGCTGCTGACAAACATGACAGAGAAATCAATTTTCAAAATTAATAA 2160
Db 2101 AAGTTGAAACAGCTGCTGACAAACATGACAGAGAAATCAATTTTCAAAATTAATAA 2160
Qy 2161 AGTTATATCAATCTGATGTTGATAGAGACTCTCTTCCCTGAAATCTGGGTTTTC 2220
Db 2161 AGTTATATCAATCTGATGTTGATAGAGACTCTCTTCCCTGAAATCTGGGTTTTC 2220
Qy 2221 AACAAATTAACACTTGAGAGACAGTAAATGAACAGTTTTC 2262
Db 2221 AACAAATTAACACTTGAGAGACAGTAAATGAACAGTTTTC 2262

RESULT 3
ABK51944
ID ABK51944 standard; cDNA; 2214 BP.
XX
AC ABK51944;
XX
DT 27-AUG-2002 (first entry)
XX
DE cDNA encoding human G-protein coupled receptor HGRBM5.
XX
KW Human; G-protein coupled receptor; GPCR; HGRBM5; colon; brain;
KW ovary; thymus; lung; immune system; cancer; immune disorder;
KW neurological disorder; infection; human immunodeficiency virus; HIV;
KW antiallergic; antiasmatic; dermatological; antiarteriosclerotic;
KW antitumor; antidiabetic; nephrotoxic; osteopathic; antiarthritic;
KW antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnery;
KW viricide; antibacterial; antifungal; antiparasitic; protozoacide;
KW antihelmintic; nootropic; neuroprotective; antidepressant;
KW anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2214
FT /tag= a
FT /product= "GPCR HGRBM5"
XX
PN MO20026824-A2.
XX
PD 04-APR-2002.
XX
PF 26-SEP-2001; 2001WO-US30365.
XX
PR 27-SEP-2000; 2000US-235713P.
PR 16-JUN-2001; 2001US-261781P.
PR 19-JUL-2001; 2001US-30605P.
PR 03-AUG-2001; 2001US-310436P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

```

XX Feder JN, Mintier G, Ramanathan CS, Hawken DR;
XX MPI: 2002-435196/46.
DR P-PSDB: AAU97158.
XX Novel G protein-coupled receptor, HGRBMV5 polypeptide, useful for
PT treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
XX osteoarthritis -
XX
PS Claim 1; Fig 1; 148bp; English.
XX The present invention relates to the isolation of a novel human
CC G-protein coupled receptor (GPCR) (HGRBMV5), and the polynucleotide
CC sequence encoding it. The HGRBMV5 polypeptide and polynucleotide
CC are useful for preventing, treating or ameliorating a disease,
CC disorder or condition related to the colon, brain, ovaries, thymus,
CC lungs or immune system. They are particularly useful for the
CC treatment or prevention of cancers, immune disorders, neurological
CC disorders, and diseases related to the brain, ovaries, thymus or
CC lungs. The polynucleotide sequence is useful for diagnosing or
CC determining susceptibility to infections such as bacterial, fungal,
CC protozoan and viral infections, particularly infections caused by
CC human immunodeficiency virus (HIV or HIV-2). The present sequence
CC encodes human GPCR HGRBMV5.
XX
SQ Sequence 2214 BP: 638 A; 457 C; 416 G; 703 T; 0 other;

Query Match 97.7%; Score 2214; DB 24; Length 2214;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 ATGTTCTTCTACTCATTTTCATGCTTGATGATGCAAAAGATTGTCATGCACTCA 111
DB 1 ATGTTCTTCTACTCATTTTCATGCTTGATGATGCAAAAGATTGTCATGCACTCA 60
OY 112 GGTAGCATGATCCTCTTCATGCCAAAAGATATTTCCCTGTGGAACTTTACCAAG 171
DB 61 GGTAGCATGATCCTCTTCATGCCAAAAGATATTTCCCTGTGGAACTTTACCAAG 120
OY 172 TGTCTTACCCCGAGCTTTTCACTGTATGAGCAAGATGACTGTGGAGAGGGCGGACGAA 231
DB 121 TGTCTTACCCCGAGCTTTTCACTGTATGAGCAAGATGACTGTGGAGAGGGCGGACGAA 180
OY 232 GAGAACTGTGATGACTAGTGTGATGGGAGCATATTTGGCAGTGGATGGAATGCT 291
DB 181 GAGAACTGTGATGACTAGTGTGATGGGAGCATATTTGGCAGTGGATGGAATGCT 240
OY 292 AACAGCGTGGCTTTAACAAGAGAGTCTTTCTAAAACAGTATCCAAATGCTGTGACTGC 351
DB 241 AACAGCGTGGCTTTAACAAGAGAGTCTTTCTAAAACAGTATCCAAATGCTGTGACTGC 300
OY 352 AAAAGAACTGAAATGTGTAAATGTGTAAAGTCTGTGTCCGATGATTTCTAAC 411
DB 301 AAAAGAACTGAAATGTGTAAATGTGTAAAGTCTGTGTCCGATGATTTCTAAC 360
OY 412 AATGTGACATTAAGTCTCTTAAGAAAAAACAATCCACAGTCTCCAGATTAAGTTTC 471
DB 361 AATGTGACATTAAGTCTCTTAAGAAAAAACAATCCACAGTCTCCAGATTAAGTTTC 420
OY 472 ATCAAAATACAAAACTTAAAAAGATATTTCTTGACATTAATGGATTAGACATATCC 531
DB 421 ATCAAAATACAAAACTTAAAAAGATATTTCTTGACATTAATGGATTAGACATATCC 480
OY 532 AGGAAGAGATTTTGGATTATGTATCTGCAAAATATATATCTCAACCAACCTGCATC 591
DB 481 AGGAAGAGATTTTGGATTATGTATCTGCAAAATATATATCTCAACCAACCTGCATC 540
OY 592 ACAACCTCAGACCTGAAATATTTCAAGACTTACATCAGCTAATCTGGCTAATTTAGAT 651
DB 541 ACAACCTCAGACCTGAAATATTTCAAGACTTACATCAGCTAATCTGGCTAATTTAGAT 600

OY 652 GACAATCCAAATTAACAGAAATTTCAAGCGCTTGTATAGGAGTAAATCTGTTTTC 711
DB 601 GACAATCCAAATTAACAGAAATTTCAAGCGCTTGTATAGGAGTAAATCTGTTTTC 660
OY 712 CTGTCTATGTTAATTAATTAAGTCTTCCCAAGCATGTGTGCCAAATGCT 771
DB 661 CTGTCTATGTTAATTAATTAAGTCTTCCCAAGCATGTGTGCCAAATGCT 720
OY 772 CAATCACTAGGGGAGATTTGAAAGGCAATGAAATTAATCTCACAAATTTACGTT 831
DB 721 CAATCACTAGGGGAGATTTGAAAGGCAATGAAATTAATCTCACAAATTTACGTT 780
OY 832 CTGTGTGCGAATTTGCTCACAGTCTGTTCTGCTGAAATCAATTTGTTTCCA 891
DB 781 CTGTGTGCGAATTTGCTCACAGTCTGTTCTGCTGAAATCAATTTGTTTCCA 840
OY 892 GAGAAACATTTTCTTCAATTAATAATTTAGAGAACTGATCTGTCTAGCAATTCGATA 951
DB 841 GAGAAACATTTTCTTCAATTAATAATTTAGAGAACTGATCTGTCTAGCAATTCGATA 900
OY 952 ACGAGCTATCAGCTCACTTTTAAAGACTTGAAGCTTCAAAAAGCTGAACCTGCA 1011
DB 901 ACGAGCTATCAGCTCACTTTTAAAGACTTGAAGCTTCAAAAAGCTGAACCTGCA 960
OY 1012 TCCAACTCTTATGATCTTCAAGAACCAAGTTGAAAGCTTTAAACAATTCACTCT 1071
DB 961 TCCAACTCTTATGATCTTCAAGAACCAAGTTGAAAGCTTTAAACAATTCACTCT 1020
OY 1072 CTAACTCTGAAAGATTAAGATTCCAATTAATAACAGAAATGTTTCAACCAATGAG 1131
DB 1021 CTAACTCTGAAAGATTAAGATTCCAATTAATAACAGAAATGTTTCAACCAATGAG 1080
OY 1132 AATCTTCTCAGATTTTCAAAAACCTTGTGATAGCTCTGTAGCTCCCATGTCGGA 1191
DB 1081 AATCTTCTCAGATTTTCAAAAACCTTGTGATAGCTCTGTAGCTCCCATGTCGGA 1140
OY 1192 ATATGATGCCCTTGACGAGAGGCAATTTCTATTTGAGAGCTCTTGCTTAACAATATC 1251
DB 1141 ATATGATGCCCTTGACGAGAGGCAATTTCTATTTGAGAGCTCTTGCTTAACAATATC 1200
OY 1252 CTGAGATATTTGTCTGAGGTATAGCTTCAATTAACGCTTGTGAAATCTTTTGTGATT 1311
DB 1201 CTGAGATATTTGTCTGAGGTATAGCTTCAATTAACGCTTGTGAAATCTTTTGTGATT 1260
OY 1312 GGCATGAGATCTTCAATTAAGCTGAATAATCACTACAGTATGTCATCAAAATCTT 1371
DB 1261 GGCATGAGATCTTCAATTAAGCTGAATAATCACTACAGTATGTCATCAAAATCTT 1320
OY 1372 TGTGTGTGATGCTCTGATGAGTGTATCTTGTCTTGTGATTTGATTAATAA 1431
DB 1321 TGTGTGTGATGCTCTGATGAGTGTATCTTGTCTTGTGATTTGATTAATAA 1380
OY 1432 TACGAGGAGCAATTAAGAGTATGCTGTGTGAGATGAGAACGTCAGTGGCGCTC 1491
DB 1381 TACGAGGAGCAATTAAGAGTATGCTGTGTGAGATGAGAACGTCAGTGGCGCTC 1440
OY 1492 ATGGGGTTCCTGGCCATGCTGTCCACCGAAAGTCTCTGTCTGCTACTGACCTTGTGACT 1551
DB 1441 ATGGGGTTCCTGGCCATGCTGTCCACCGAAAGTCTCTGTCTGCTACTGACCTTGTGACT 1500
OY 1552 TTGAGAAAGTTCCTGGTCATTTCTTCCCTTCAAGTAACTTGAACCTGAAAAAGCGAG 1611
DB 1501 TTGAGAAAGTTCCTGGTCATTTCTTCCCTTCAAGTAACTTGAACCTGAAAAAGCGAG 1560
OY 1612 ACCTCAGTATCCATTTGATGATGATGAGCGGAGTTTAAATAGCTGTAATCCATTT 1671
DB 1561 ACCTCAGTATCCATTTGATGATGATGAGCGGAGTTTAAATAGCTGTAATCCATTT 1620
OY 1672 TGAATTAAGATTTTGTGAAACCTTTATAGGAAAAATGAGATGTTTCCACTTAT 1731
DB 1621 TGAATTAAGATTTTGTGAAACCTTTATAGGAAAAATGAGATGTTTCCACTTAT 1680
OY 1732 TATGACCAACAGAAATATTTGAGAGCAAGGATATTTCTTGGAAATTTTCTTAGGTG 1791

Db 1681 TATGACCAACAGAGATATGGAGGAAAGGATTCCTTGGAAATTTCCATAGGTG 1740
Qy 1792 AACTTGTGCTTTTCTCATCATTTGTGTTTCCATATATATATGTTCTGTTCCATTCAA 1851
Db 1741 AACTTGTGCTTTTCTCATCATTTGTGTTTCCATATATATATGTTCTGTTCCATTCAA 1800
Qy 1852 AAAACGCGCTTGGAGACACAGAGTAAGAAATTTGTTTGGAGAGAGGTGGCTGTGCA 1911
Db 1801 AAAACGCGCTTGGAGACACAGAGTAAGAAATTTGTTTGGAGAGAGGTGGCTGTGCA 1860
Qy 1912 AATCGTTTCTTTTATATAGTGTCTCTGATGCCATCTGCTGATTCCTGTATTTGTAATT 1971
Db 1861 AATCGTTTCTTTTATATAGTGTCTCTGATGCCATCTGCTGATTCCTGTATTTGTAATT 1920
Qy 1972 AAAATCTTTTCTTCTTCCGGGTGGAAATACACAGACCAATGACTTCTGGATATGTAAT 2031
Db 1921 AAAATCTTTTCTTCTTCCGGGTGGAAATACACAGACCAATGACTTCTGGATATGTAAT 1980
Qy 2032 TTTTTCCTTCCAGTTACAGAGCTTGTGATCCATCTCTATATCTCTACACCAACTTT 2091
Db 1981 TTTTTCCTTCCAGTTACAGAGCTTGTGATCCATCTCTATATCTCTACACCAACTTT 2040
Qy 2092 TTTAAGACAGAGTTGAAACAGCTGCTGACAAACATCAGAGAAATCAATTTCAAAAT 2151
Db 2041 TTTAAGACAGAGTTGAAACAGCTGCTGACAAACATCAGAGAAATCAATTTCAAAAT 2100
Qy 2152 AAAAAAAGTTTAT 2211
Db 2101 AAAAAAAGTTTAT 2160
Qy 2212 GGGGTTTGAACAATAATACATCTGAGACAGTATATGAACCAAGTTTCTGAG 2265
Db 2161 GGGGTTTGAACAATAATACATCTGAGACAGTATATGAACCAAGTTTCTGAG 2214

RESULT 4
ID ABL40197 standard; cDNA; 2190 BP.
XX
AC ABL40197;
XX
DT 23-MAY-2002 (first entry)
XX
DE Human G protein-coupled receptor TGR17-6 encoding cDNA SEQ ID NO:16.
XX
KW Human; G protein-coupled receptor; TGR17-6; nootropic; antiinflammatory;
KW vasotrophic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW neurological; inflammatory; circulatory; degenerative; immune system;
KW digestive disease; cancer; gene; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..2190
FT /*Lag= a
FT /partial
FT /product= "TGR17-6"
FT /note= "no stop codon given"
XX
XX
PN W0200204640-A1.
XX
XX 17-JAN-2002.
XX
XX 06-JUL-2001; 2001WO-JP05878.
XX
XX 07-JUL-2000; 2000JP-0211989.
XX
XX 18-DEC-2000; 2000JP-0383794.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX PA
XX PI Moriya T, Ito T, Shintani Y, Miyajima N;

DR WP1: 2002-179706/23.
DR P-PSDB; ABB06255.
XX
XX G-protein coupled receptor protein TGR17 of human origin and DNA
PT encoding it for diagnosis and treatment of cancer and circulatory and
PT other diseases associated with its expression -
XX
XX
PS Claim 7; Page 135-136; 145pp; Japanese.
XX
XX The present invention describes a human guanine nucleotide binding
CC protein (G protein) coupled receptor protein designated TGR17, which has
CC 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC antiinflammatory, vasotrophic, immunomodulator and cytostatic activities.
CC The TGR17 polynucleotides and protein can be used in gene therapy and
CC protein therapy. G-protein coupled receptor proteins are cell membrane
CC proteins mediating the cellular response to a large variety of signaling
CC molecules. The TGR17 polynucleotides and proteins can be used in the
CC diagnosis, treatment and prevention of diseases including neurological,
CC inflammatory, circulatory, degenerative, immune system and digestive
CC diseases and cancer. The present sequence encodes human TGR17-6 from the
CC present invention.
XX
SQ Sequence 2190 BP; 626 A; 457 C; 409 G; 698 T; 0 other;
XX
Query Match 92.9%; Score 2104.8; DB 24; Length 2190;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 2188; Conservative 0; Mismatches 2; Indels 72; Gaps 1;
XX
Qy 1 ATGATGTTTTTCTGTTTTTAAACATCTCTTCAAGCTCAGATTGATTAATATTTCTTT 60
Db 1 ATGATGTTTTTCTGTTTTTAAACATCTCTTCAAGCTCAGATTGATTAATATTTCTTT 60
Qy 61 CTACTCATTTTCACTGTTCTGATCAATGTCAAAGATTGTCATGCTCAAGGTACATG 120
Db 61 CTACTCATTTTCACTGTTCTGATCAATGTCAAAGATTGTCATGCTCAAGGTACATG 120
Qy 121 ATCACTCTTTCATGCGCAAAAGATATTTTCCCTGTGGGAATCTTTCACAAAGTCTTACC 180
Db 121 ATCACTCTTTCATGCGCAAAAGATATTTTCCCTGTGGGAATCTTTCACAAAGTCTTACC 180
Qy 181 CGAGCTTTTCACTGTGATGCAAGATGACCTGTGGACCGGGCGGACGAAGAACTGT 240
Db 181 CGAGCTTTTCACTGTGATGCAAGATGACCTGTGGACCGGGCGGACGAAGAACTGT 240
Qy 241 GGTGACATGATGATGAGGCGACCATTTTGGACAGTGCATGGAATGCTTAAACAGCTG 300
Db 241 GGTGACATGATGATGAGGCGACCATTTTGGACAGTGCATGGAATGCTTAAACAGCTG 300
Qy 301 GCCTTAACACAGAGAGCTTTCTTAAACAGTATCCCAATGCTGTGACTGCAAGAAACT 360
Db 301 GCCTTAACACAGAGAGCTTTCTTAAACAGTATCCCAATGCTGTGACTGCAAGAAACT 360
Qy 361 GAATTGGAATGTGTAAATGATGATTAAGCTGTGTCGATGATTTCTTAAAGTGTGACA 420
Db 361 GAATTGGAATGTGTAAATGATGATTAAGCTGTGTCGATGATTTCTTAAAGTGTGACA 420
Qy 421 TTACTGTCTTAAAGAAAACAAATTCACAGTCTTCCAGATTAAGTATTTCATCAATAC 480
Db 421 TTACTGTCTTAAAGAAAACAAATTCACAGTCTTCCAGATTAAGTATTTCATCAATAC 480
Qy 481 ACAAAACCTTAAAGATATTTCTTCAAGATTAATGATTAAGACATATCCAGAAAAGA 540
Db 481 ACAAAACCTTAAAGATATTTCTTCAAGATTAATGATTAAGACATATCCAGAAAAGA 540
Qy 541 TTTTGTGATATGTAATCTGCAAAATATATATCTCAACCAACATGCAACACCTTC 600
Db 541 TTTTGTGATATGTAATCTGCAAAATATATATCTCAACCAACATGCAACACCTTC 600
Qy 601 AGACTGGAATATTTCAAGACTTACATCAGTACTTGGCTAATTTCTAATGACATCCA 660
Db 601 AGACTGGAATATTTCAAGACTTACATCAGTACTTGGCTAATTTCTAATGACATCCA 660

Oy	661	ATAACGAAATTTCAACGCGCTGTGTTTACGGGATTTAAATTCCTGTTTCCGTCTATG	720
Dp	661	ATAACGAAATTTCAACGCGCTGTGTTTACGGGATTTAAATTCCTGTTTCCGTCTATG	720
Oy	721	GTTAATACTACTAGAACTCTTCCAGCAGATGTGTGCCAAATGCTCACTCAAC	780
Dp	721	GTTAATACTACTAGAACTCTTCCAGCAGATGTGTGCCAAATGCTCACTCAAC	780
Oy	781	TGGGTGATTTTGGGAAGGCAATAGAAATCAATCTCAAAATTTACGTTTCTGTCTGC	840
Dp	781	TGGGTGATTTTGGGAAGGCAATAGAAATCAATCTCAAAATTTACGTTTCTGTCTGC	840
Oy	841	GATTGCGCTCAGACGTCTGTCTCTGCTAGAAATCAAAATGGTTTGTTCAGAGAGACAA	900
Dp	841	GATTGCGCTCAGACGTCTGTCTCTGCTAGAAATCAAAATGGTTTGTTCAGAGAGACAA	900
Oy	901	TTTTCTTCTATTTAAAAAATTTAGAGAACTGGAATCTGTCTAGCAATACGATACGAGCTA	960
Dp	856	-----CTGATCTGTCTAGCAATACGATACGAGCTA	888
Oy	961	TCACCTCACTTTTAAAAAGCTTGAAGCTTCAAAAAAGCTGAACCTGTCACTCAACT	1020
Dp	889	TCACCTCACTTTTAAAAAGCTTGAAGCTTCAAAAAAGCTGAACCTGTCACTCAACT	948
Oy	1021	CTTATGTATCTTCACAAGAACCAAGTTGAAAGCTTAAACAATCTCAAGTCTTAGACCTG	1080
Dp	949	CTTATGTATCTTCACAAGAACCAAGTTGAAAGCTTAAACAATCTCAAGTCTTAGACCTG	1008
Oy	1081	GAAAGGATAGAGATTCCAAATATATAACAAGAAATGTTTCAACCCATGAAAGAACTTCT	1140
Dp	1009	GAAAGGATAGAGATTCCAAATATATAACAAGAAATGTTTCAACCCATGAAAGAACTTCT	1068
Oy	1141	CACATTTATTTCAAAAACTTTCGATACGTCTCATGTCTCCCAATGTCCGAATATGTATG	1200
Dp	1069	CACATTTATTTCAAAAACTTTCGATACGTCTCATGTCTCCCAATGTCCGAATATGTATG	1128
Oy	1201	CCCTTGAACGACCGCATTTCTTCATTGAGAGACCTCTTGAGTAAACAATATCTTCAGAAATA	1260
Dp	1129	CCCTTGAACGACCGCATTTCTTCATTGAGAGACCTCTTGAGTAAACAATATCTTCAGAAATA	1188
Oy	1261	TTTGTCTGGGTTATAGCTTTCATTAAACGCTTTGGAAGATCTTTTGTGATTTGGCATGAGA	1320
Dp	1189	TTTGTCTGGGTTATAGCTTTCATTAAACGCTTTGGAAGATCTTTTGTGATTTGGCATGAGA	1248
Oy	1321	TCCTTCATTTAAAGCTGAAATAACACTCAGCTATGTCCATCAAAATCTTGTGTGCT	1380
Dp	1249	TCCTTCATTTAAAGCTGAAATAACACTCAGCTATGTCCATCAAAATCTTGTGTGCT	1308
Oy	1381	GATTGCCGTAAGGGGTGTTACTGTCTCTTGTGGAATTTTGGATATATAATACCGAGGG	1440
Dp	1309	GATTGCCGTAAGGGGTGTTACTGTCTCTTGTGGAATTTTGGATATATAATACCGAGGG	1368
Oy	1441	CAGTATCAGAAAGTATGCCCTTGCCTGTGTGATGAGAGACGTCAGTGCCTGATGGGGTTC	1500
Dp	1369	CAGTATCAGAAAGTATGCCCTTGCCTGTGTGATGAGAGACGTCAGTGCCTGATGGGGTTC	1428
Oy	1501	CTGGGCATGCTGTCCACCGAAGTCTCTGTCTGTCTACTAGACTTACTTGTGGAAGAG	1560
Dp	1429	CTGGGCATGCTGTCCACCGAAGTCTCTGTCTGTCTACTAGACTTACTTGTGGAAGAG	1488
Oy	1561	TTCCCTGATCATGTCTTCCCTTCAGTAAACCTTGCACCTGGAAAAACGACAGCTCAAGTC	1620
Dp	1489	TTCCCTGATCATGTCTTCCCTTCAGTAAACCTTGCACCTGGAAAAACGACAGCTCAAGTC	1548
Oy	1621	ATCTCATTTGCAATCTGATGAGCGGGATTTTAAATAGCTGTAAATTCATTTTGGAAATAG	1680
Dp	1549	ATCTCATTTGCAATCTGATGAGCGGGATTTTAAATAGCTGTAAATTCATTTTGGAAATAG	1608
Oy	1681	GATTATTTTGGAAAACTTTTATGAGGAAAAATGAGAGATGTTCCCACTTATATTAAGCAAA	1740
Dp	1609	GATTATTTTGGAAAACTTTTATGAGGAAAAATGAGAGATGTTCCCACTTATATTAAGCAAA	1668
Oy	1741	ACAGAGATATTTGGAACAAAGGATATCTCTTGGAAATTTTCTAGATGTGAACTTGTCTG	1800

Db	1669	ACAGAAAGTATTGGAGCAAGGATATCTCTTGGATTTTCCAGGTGGAATTCTG	1728
Qy	1801	GCTTTTCATCATGTGTGTTTTCTATATATTAATAATGTTCCTCATTTCAAAAAACGCC	1860
Db	1729	GCTTTTCATCATGTGTGTTTTCTATATTAATAATGTTCCTCATTTCAAAAAACGCC	1788
Qy	1861	TTGGAGACCAAGAAAGTAAGGAATGTGTTTGGAAAGAGGGGCTGTGCAATCGTTCC	1920
Db	1789	TTGGAGACCAAGAAAGTAAGGAATGTGTTTGGAAAGAGGGGCTGTGCAATCGTTTC	1848
Qy	1921	TTTTTTATAGTGTCTCTGATGCCATCTGCTGATTCCTGTATTTGTATTAATCCTT	1980
Db	1849	TTTTTTATAGTGTCTCTGATGCCATCTGCTGATTCCTGTATTTGTATTAATCCTT	1908
Qy	1991	TCCCTCTTCCGGGTGGAAATPACAGACAAATGAATCTTCCGGAAAGATTTTTCCTT	2040
Db	1909	TCCCTCTTCCGGGTGGAAATPACAGACAAATGAATCTTCCGGAAAGATTTTTCCTT	1968
Qy	2041	CCAGTTAAACAGTGTGTAATCCAAATCCTCTATCTCTCAACCAACTTTTTTAAGAC	2100
Db	1969	CCAGTTAAACAGTGTGTAATCCAAATCCTCTATCTCTCAACCAACTTTTTTAAGAC	2028
Qy	2101	AAGTTGAAACAGCTGCTGCACAAACATCAGAGAAATCATTTCAAAATTTAAAAAAA	2160
Db	2029	AAGTTGAAACAGCTGCTGCACAAACATCAGAGAAATCATTTCAAAATTTAAAAAAA	2088
Qy	2161	AGTTTATCATCATCTTGTGTGATGAGGACCTCTTCCCGGAAACTGTGGGTTTG	2220
Db	2089	AGTTTATCATCATCTTGTGTGATGAGGACCTCTTCCCGGAAACTGTGGGTTTG	2148
Qy	2221	AACAAATATACACTTGGAGACAGTATATATGAAACAGTTTCC	2262
Db	2149	AACAAATATACACTTGGAGACAGTATATGAAACAGTTTCC	2190
RESULT 5			
ABK51947			
ID	ABK51947	standard; cDNA; 2142 BP.	
XX	AC		
XX	ABK51947;		
XX	27-AUG-2002	(first entry)	
DT			
DE	cDNA encoding human G-protein coupled receptor HGRBM5 splice variant.		
XX			
XX			
KW	Human; G-protein coupled receptor; GPCR; HGRBM5; colon; brain;		
KW	ovary; lymphus; lung; immune system; cancer; immune disorder;		
KW	neurological disorder; infection; human immunodeficiency virus; HIV;		
KW	antiallergic; antiaesthetic; dermatological; antiarteriosclerotic;		
KW	antiviral; antididiabetic; nephrotoxic; osteopathic; antiarthritic;		
KW	antiinflammatory; antirheumatic; antichyroid; cytostatic; vulnerary;		
KW	vincinide; antibacterial; antifungal; antiparasitic; protozoacide;		
KW	antihelmintic; nootropic; neuroprotective; antidepressant;		
XX	anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..2142	
FT		/*tag= a	
FT		/product= "GPCR HGRBM5"	
XX			
XX	WO200226824-A2.		
XX			
PD	04-APR-2002.		
XX			
PF	26-SEP-2001; 2001WO-US30365.		
XX			
PR	27-SEP-2000; 2000US-235713P.		
PR	16-JAN-2001; 2001US-261781P.		
PR	19-JUL-2001; 2001US-306605P.		
PR	03-AUG-2001; 2001US-310436P.		

XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Feder JN, Mintier G, Ramanathan CS, Hawken DR;
XX WPI; 2002-435196/46.
XX P-PSDB; AAU97159.
XX
XX Novel G protein-coupled receptor, HGPBMY5 polypeptide, useful for
XX treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
XX colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
XX osteoarthritis -
XX
XX Claim 1; Fig 5; 148pp; English.
XX
XX The present invention relates to the isolation of a novel human
XX G-protein coupled receptor (GPCR) (HGPBMY5), and the polynucleotide
XX sequence encoding it. The HGPBMY5 polypeptide and polynucleotide
XX are useful for preventing, treating or ameliorating a disease,
XX disorder or condition related to the colon, brain, ovaries, thymus,
XX lungs or immune system. They are particularly useful for the
XX treatment or prevention of cancers, immune disorders, neurological
XX disorders, and diseases related to the brain, ovaries, thymus or
XX lungs. The polynucleotide sequence is useful for diagnosing or
XX determining susceptibility to infections such as bacterial, fungal,
XX protozoan and viral infections, particularly infections caused by
XX human immunodeficiency virus (HIV or HIV-2). The present sequence
XX encodes a splice variant of human GPCR HGPBMY5.
XX
XX Sequence 2142 BP; 613 A; 435 C; 408 G; 686 T; 0 other;
XX
XX Query Match 90.9%; Score 2060; DB 24; Length 2142;
XX Best Local Similarity 96.7%; Pred. No. 0;
XX Matches 2142; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
XX
QY 52 ATGTTCTTCTACTTTCATTCATCGTTCTGATCAATGTCAAAGATTTTGACACTGACTCAA 111
DB 1 ATGTTCTTCTACTTTCATTCATCGTTCTGATCAATGTCAAAGATTTTGACACTGACTCAA 60
QY 112 GGTAGCATGATCACTCTTTCATGCGCAAAAGAGATTTTCCCTGTGGGAATCTTACCAG 171
DB 61 GGTAGCATGATCACTCTTTCATGCGCAAAAGAGATTTTCCCTGTGGGAATCTTACCAG 120
QY 172 TGCTTACCCGAGCTTTTCACTGTGATGGCAAGGATGACTGTGGGAACGGGGCGAGCAAA 231
DB 121 TGCTTACCCGAGCTTTTCACTGTGATGGCAAGGATGACTGTGGGAACGGGGCGAGCAAA 180
QY 232 GAGAACTGTGGTGCACACTAGTGAATGGGCGACATATTTGGCACAAGTGAATAGCT 291
DB 181 GAGAACTGTGGTGCACACTAGTGAATGGGCGACATATTTGGCACAAGTGAATAGCT 240
QY 292 AACGCGGTGGCTTAAACACAGAGATGCTTTTAAACACATTCACAAATGCTGTGATGC 351
DB 241 AACGCGGTGGCTTAAACACAGAGATGCTTTTAAACACATTCACAAATGCTGTGATGC 300
QY 352 AAAGAACTGAATGGAATGTAAATGGTGAATTAAGTCTGACGATGATTTCTAAC 411
DB 301 AAAGAACTGAATGGAATGTAAATGGTGAATTAAGTCTGACGATGATTTCTAAC 360
QY 412 AATGTGACATTAAGTCTGTTAAGAAAAACAAATTCACAGTCTTCCAGATTAAGTTTC 471
DB 361 AATGTGACATTAAGTCTGTTAAGAAAAACAAATTCACAGTCTTCCAGATTAAGTTTC 420
QY 472 ATCAAAATACCAAAACCTTAAAGATATTTCTTACAGATAATTCGATTAACATATTC 531
DB 421 ATCAAAATACCAAAACCTTAAAGATATTTCTTACAGATAATTCGATTAACATATTC 480
QY 532 AGGAAGCATTTTGGATTATGTAACTGCAATATTAATCTCAACCAACATGATC 591
DB 481 AGGAAGCATTTTGGATTATGTAACTGCAATATTAATCTCAACCAACATGATC 517
QY 592 ACAACCTCGACCTGGAAATATTCAAAAGATTACATGACTTAAGTCTTAATCTAGAT 651
|||||

DB 518 -----TAACTCAGAT 528
QY 652 GACAAATCCAAATACCAAGATTTTCAAGCGCTGTGTTACGGGATTAATTCCTGTTTTC 711
DB 529 GACAAATCCAAATACCAAGATTTTCAAGCGCTGTGTTACGGGATTAATTCCTGTTTTC 588
QY 712 CTGTCTATGGTTAATTAATCTAGTAAGCTCTTCCCAAGCAGATGTGTCGCAATAGCTT 771
DB 589 CTGTCTATGGTTAATTAATCTAGTAAGCTCTTCCCAAGCAGATGTGTCGCAATAGCTT 648
QY 772 CAACTCAACTGGGTGATTTGGAAAGCAATAGAAATTAATCTCAAAATTTCTAGTT 831
DB 649 CAACTCAACTGGGTGATTTGGAAAGCAATAGAAATTAATCTCAAAATTTCTAGTT 708
QY 832 CTGTCTATGGTTAATTAATCTAGTAAGCTCTTCCCAAGCAGATGTGTCGCAATAGCTT 891
DB 709 CTGTCTATGGTTAATTAATCTAGTAAGCTCTTCCCAAGCAGATGTGTCGCAATAGCTT 768
QY 892 GAGAAACATTTTCTTCAATTAAGAAATTTAGAGAACTGGATCTGTCTAGCAATAGATTA 951
DB 769 GAGAAACATTTTCTTCAATTAAGAAATTTAGAGAACTGGATCTGTCTAGCAATAGATTA 828
QY 952 ACGAGACTATCACTTCACTTTTAAAGACTTGAAGCTTCTACAAAGCTGAACTGTCA 1011
DB 829 ACGAGACTATCACTTCACTTTTAAAGACTTGAAGCTTCTACAAAGCTGAACTGTCA 888
QY 1012 TCCAAATCTCTTATGTATCTTCAAGAAACCAAGTTGAAAGCTTAAACAATTCAGTCT 1071
DB 889 TCCAAATCTCTTATGTATCTTCAAGAAACCAAGTTGAAAGCTTAAACAATTCAGTCT 948
QY 1072 CTAGACCTGGAAGAGATAGAGATTCGAATATTAACACAGATGTTTCAACCCATGAAG 1131
DB 949 CTAGACCTGGAAGAGATAGAGATTCGAATATTAACACAGATGTTTCAACCCATGAAG 1008
QY 1132 AATCTTCTCACTTATTTTCAAAAACTTTCGATCTGCTATGCTCCCATGTGCCA 1191
DB 1009 AATCTTCTCACTTATTTTCAAAAACTTTCGATCTGCTATGCTCCCATGTGCCA 1068
QY 1192 ATATGTATGCCCTTGAACGAGCAATTTCTCAATTTGAAGACCTCTTGCTAAACATATTC 1251
DB 1069 ATATGTATGCCCTTGAACGAGCAATTTCTCAATTTGAAGACCTCTTGCTAAACATATTC 1128
QY 1252 CTCGAATATTTGTCTGAGTTATAGCTTCAATTAAGCTTGAAGATCTTTTGTCAAT 1311
DB 1129 CTCGAATATTTGTCTGAGTTATAGCTTCAATTAAGCTTGAAGATCTTTTGTCAAT 1188
QY 1312 GGCATGAGATCTTCAATTAAGCTGAAATTAACAATCAAGCTATGTCCATCAAAATCTT 1371
DB 1189 GGCATGAGATCTTCAATTAAGCTGAAATTAACAATCAAGCTATGTCCATCAAAATCTT 1248
QY 1372 TGTGTGCTGATTTGCTGAGTGGTGTGTTTCTGTTTGGCATTTTGGATTAATAA 1431
DB 1249 TGTGTGCTGATTTGCTGAGTGGTGTGTTTCTGTTTGGCATTTTGGATTAATAA 1308
QY 1432 TACCGAGGAGATGATGAGATGATGCTTGTGTGATGATGAGAGAGCTGAGTGCCTC 1491
DB 1309 TACCGAGGAGATGATGAGATGATGCTTGTGTGATGATGAGAGAGCTGAGTGCCTC 1368
QY 1492 ATGGGCTTCTGAGCATGCTGTTCACAGAAATGCTCTGTCTGCTATGACATTAAGT 1551
DB 1369 ATGGGCTTCTGAGCATGCTGTTCACAGAAATGCTCTGTCTGCTATGACATTAAGT 1428
QY 1552 TTGGAAGAGTTCTGTGATGATGCTTCCCTTCAAGTAACTTGAACCTGGAAGAGCGG 1611
DB 1429 TTGGAAGAGTTCTGTGATGATGCTTCCCTTCAAGTAACTTGAACCTGGAAGAGCGG 1488
QY 1612 ACCTCAGTCAATCTTCAATTTGATGATGAGGAGGATTTTAAATAGCTGATTAATTCAT 1671
DB 1489 ACCTCAGTCAATCTTCAATTTGATGATGAGGAGGATTTTAAATAGCTGATTAATTCAT 1548
QY 1672 TGGAAATTAAGATTTTGGAAACTTTTATGGAAGAAATGAGATAGTTTCCACTTAT 1731
DB 1549 TGGAAATTAAGATTTTGGAAACTTTTATGGAAGAAATGAGATAGTTTCCACTTAT 1608
|||||

QY 1732 TATGACCAACAGAGATATTGGAAGCAAGGATTTCTTGAATTTTCTAGGTG 1791
 Db 1609 TATGACCAACAGAGATATTGGAAGCAAGGATTTCTTGAATTTTCTAGGTG 1668
 QY 1792 AACTGTGCTTTTCTCATCTGTGTGTTTCTATATTAATTAATGTTCTGTTCAATCA 1851
 Db 1669 AACTGTGCTTTTCTCATCTGTGTGTTTCTATATTAATTAATGTTCTGTTCAATCA 1728
 QY 1852 AAAACCGCTTGCAGACCAAGAGTAAGTAATGTTTGAAGAGAGGTGCTGTGCA 1911
 Db 1729 AAAACCGCTTGCAGACCAAGAGTAAGTAATGTTTGAAGAGAGGTGCTGTGCA 1788
 QY 1912 AATGTTTCTTTTATAGTGTCTGTATGCAATCTGCTGTAATCTGTAATTTGTAAT 1971
 Db 1789 AATGTTTCTTTTATAGTGTCTGTATGCAATCTGCTGTAATCTGTAATTTGTAAT 1848
 QY 1972 AAAATCTTTTCTTCTTCCGGGTGAAATACCAAGACAAATGACTTCTGTGATGTAAT 2031
 Db 1849 AAAATCTTTTCTTCTTCCGGGTGAAATACCAAGACAAATGACTTCTGTGATGTAAT 1908
 QY 2032 TTTTCTTCTTCAAGTAAGAGTCTTGAATCCATCTCTTATCTTCAACCAACTTT 2091
 Db 1909 TTTTCTTCTTCAAGTAAGAGTCTTGAATCCATCTCTTATCTTCAACCAACTTT 1968
 QY 2092 TTTAAGCAAGTGAACAGCTGTGCAAAACATCAGAGGAATCAATTTCAAAAT 2151
 Db 1969 TTTAAGCAAGTGAACAGCTGTGCAAAACATCAGAGGAATCAATTTCAAAAT 2028
 QY 2152 AAAAAAAGTTTATCTACATCATCTGTGTGATGAGAGACTCTCTTCCCTGAACTT 2211
 Db 2029 AAAAAAAGTTTATCTACATCATCTGTGTGATGAGAGACTCTCTTCCCTGAACTT 2088
 QY 2212 GGGTTTGAACAAATAAAGCTTGAACAGTATTAATGAACGATTTCCCTAG 2265
 Db 2089 GGGTTTGAACAAATAAAGCTTGAACAGTATTAATGAACGATTTCCCTAG 2142
 RESULT 6
 AAF88580
 ID AAF88580 standard; cDNA; 2181 BP.
 AC AAF88580;
 XX
 DT 19-NOV-2002 (first entry)
 XX
 DE Human GCRC-1 cDNA INCYTE ID 7485090CD1 SEQ ID 49.
 XX
 KW GCRC; Human; G-protein coupled receptor; anti-HIV; antiarteriosclerotic;
 KW cytosolic; neuroprotective; antiparkinsonian; hepatocrotic; laxative;
 KW cerebroprotective; antiinflammatory; virucide; antibacterial; fungicide;
 KW procoagulant; cirrhosis; cancer; stroke; Alzheimer's disease; AIDS;
 KW Parkinson's disease; Crohn's disease; constipation; infection;
 KW gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200263004-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 06-FEB-2002; 2002WO-US03635.
 XX
 PR 07-FEB-2001; 2001US-267322P.
 PR 23-FEB-2001; 2001US-271215P.
 PR 08-MAR-2001; 2001US-274551P.
 PR 23-MAR-2001; 2001US-278507P.
 PR 30-MAR-2001; 2001US-280597P.
 PR 02-APR-2001; 2001US-281107P.
 PR 06-APR-2001; 2001US-282121P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX

PI Baughn MR, Tribouley CM, Nguyen DB, Thornton M, Yao MG;
 PI Kallik DA, Gandhi AR, Walla NK, Arvizu C, Elliott VS;
 PI Harelia A, Ramkumar J, Pei J, Tang YT, Yue H, Reddy R;
 PI Buford N, Lu DM, Gaul RC, Khan FA, Walsh RT, Ison CH;
 PI Richardson TW, Griffin JA, Warren BA, Yang J, Lee EA, Harland L;
 DR WPI: 2002-627557/67.
 DR P-PSDB: AAB71322.
 XX
 PT New human G-protein coupled receptors (GCRC), useful for diagnosing or
 PT treating a disease or condition associated with decreased expression or
 PT over expression of functional GCRCs e.g. cancer, Alzheimer's and
 PT Parkinson's -
 XX
 PS Claim 110; Page 209-210; 239pp; English.

CC This invention describes novel polypeptides which have anti-HIV,
 CC antiarteriosclerotic, cytosolic, neuroprotective, antiparkinsonian,
 CC hepatocrotic, laxative, cerebroprotective, antiinflammatory, virucide,
 CC antibacterial, fungicide and procoagulant activity. The products of the
 CC invention are useful for treating a disease or condition associated with
 CC decreased expression or over expression of functional G-protein coupled
 CC receptors (GCRC), while antibodies generated against the polypeptide of
 CC the invention are useful for diagnosing a condition or disease associated
 CC with the expression of GCRC e.g. arteriosclerosis, cirrhosis, cancer,
 CC stroke, Alzheimer's disease, Parkinson's disease, Crohn's disease,
 CC constipation, AIDS, or bacterial, viral, fungal or protozoal infections.
 CC The compound described in the invention can be used for gene therapy.
 CC AAF88580-AAF88627 encode the GCRC proteins represented by
 CC AAB71322-AAF871369, described in the disclosure of the invention.
 XX
 SQ Sequence 2181 BP; 635 A; 451 C; 419 G; 676 T; 0 other;

Query Match 88.4%; Score 2002.2; DB 24; Length 2181;
 Best Local Similarity 95.4%; Pred. No. 0;
 Matches 2114; Conservative 0; Mismatches 23; Indels 78; Gaps 2;

QY 57 CTTTCTACTTCAATTCATCGTTCTGATCAATGTCAGAAATTTTGCAGTCAAGGTAG 116
 Db 39 CATGGAATTTAGTGTCAAAATGATGACAGTATGATTTTGCAGTCAAGGTAG 98
 QY 117 CATGATCACTCTTCATGCAAAAGATATTTTCCGTGGAATCTTACCAAGTCTT 176
 Db 99 CATGATCACTCTTCATGCAAAAGATATTTTCCGTGGAATCTTACCAAGTCTT 158
 QY 177 ACCCGAGCTTTTACGTGATGAGCAAGATGACGTGGAACGGGGGCGAGCAAGAA 236
 Db 159 ACCCGAGCTTTTACGTGATGAGCAAGATGACGTGGAACGGGGGCGAGCAAGAA 218
 QY 237 CTGTGTGACACTAGTGTGATGAGCAAGATATTTTGGCACTGATGAGAAATGCTAACG 296
 Db 219 CTGTGTGACACTAGTGTGATGAGCAAGATATTTTGGCACTGATGAGAAATGCTAACG 278
 QY 297 CGTGCCCTTAACACAGAGTCTTTCTAACAAGTATCCAAATGCTGTGACTGAAAGA 356
 Db 279 CGTGCCCTTAACACAGAGTCTTTCTAACAAGTATCCAAATGCTGTGACTGAAAGA 338
 QY 357 AACTGAATTTGAATGTGTAATGAGTAAAGTCTTAAAGTGTGCGAGTATTTCTAACATGT 416
 Db 339 AACTGAATTTGAATGTGTAATGAGTAAAGTCTTAAAGTGTGCGAGTATTTCTAACATGT 398
 QY 417 GACATTAATCTCTCTTAAGAAACAAATATCCAGTCTTCAAGTAAAGTCTTCAATCA 476
 Db 399 GACATTAATCTCTCTTAAGAAACAAATATCCAGTCTTCAAGTAAAGTCTTCAATCA 458
 QY 477 ATACACAAACTTAAAGATATTTTCTGACATTAATGCTATGACACATATCCAGAA 536
 Db 459 ATACACAAACTTAAAGATATTTTCTGACATTAATGCTATGACACATATCCAGAA 518
 QY 537 AGCATTTTGTGATTAATGTAATGCAAAATTTATCTGACACCAAGTCAATCAGCAAC 596
 Db 519 AGCATTTTGTGATTAATGTAATGCAAAATTTATCTGACACCAAGTCAATCAGCAAC 578

QY 597 CCTCAGACCTGGAATTTTCAAGAATTACATGACCTAATGGCTAATTCATGATGACA 656
 DB 579 CCTCAGACCTGGAATTTTCAAGAATTACATGACCTAATGGCTAATTCATGATGACA 638
 QY 657 TCCCAATTAACAGAAATTTTCAAGCGCTGTTTACGGGATTAATTCCTGTTTTCTGTC 716
 DB 639 TCCCAATTAACAGAAATTTTCAAGCGCTGTTTACGGGATTAATTCCTGTTTTCTGTC 698
 QY 717 TATGCTTAATACTACTTGAAGCTTTCCCAAGCAGATGTCGCCAAATGCTCAACT 776
 DB 699 TATGCTTAATACTACTTGAAGCTTTCCCAAGCAGATGTCGCCAAATGCTCAACT 758
 QY 777 CAACCTGGGATGATTTGGAAAGGCAATGAATTAAGTATCTCAAAATTCAGCTTCTGTC 836
 DB 759 CAACCTGGGATGATTTGGAAAGGCAATGAATTAAGTATCTCAAAATTCAGCTTCTGTC 818
 QY 837 GTGCGATTCCTCAGACGCTGTTTCTGCTAGAAATCAAAATGATTTGTTTCCAGAGAA 896
 DB 819 GTGCGATTCCTCAGACGCTGTTTCTGCTAGAAATCAAAATGATTTGTTTCCAGAGAA 837
 QY 897 GACATTTCTTCAATTAATAAATTTAGAGAACTGATCTGCTAGCAATGATTAACGGA 956
 DB 838 -----CTGATCTGCTAGCAATGATTAACGGA 866
 QY 957 GCTATCAGCTCACTTTTAAAGCTTGAAGCTTCTAACAAGCTGAACCTGCTCAACA 1016
 DB 867 ACTATACCTCACTTTTAAAGCTTGAAGCTTCTAACAAGCTGAACCTGCTCAACA 926
 QY 1017 TCTCTTATGATATCTTCAAGAACCAAGTTGAAAGCTTAAACAATTCAGTCTCTAGA 1076
 DB 927 TCTCTTATGATATCTTCAAGAACCAAGTTGAAAGCTTAAACAATTCAGTCTCTAGA 986
 QY 1077 CCTGGAAGAGATGAGATTCCAATTAATAACAGAGATGTTCAACCCATGAAGATCT 1136
 DB 987 CCTGGAAGAGATGAGATTCCAATTAATAACAGAGATGTTCAACCCATGAAGATCT 1046
 QY 1137 TTCTCAGAT-----TTATTTCAAAAACCTTGATCTGCTCTATGCTCCCATGTCG 1190
 DB 1047 TTCTCAGATCCTGTTATTTTCAAAAACCTTGATCTGCTCTATGCTCCCATGTCG 1106
 QY 1191 AATATGATGCTCTTGAACGAGCGGCAATTTCTTCAATTTAGAGACCTTGGCTAACAATAT 1250
 DB 1107 AATATGATGCTCTTGAACGAGCGGCAATTTCTTCAATTTAGAGACCTTGGCTAACAATAT 1166
 QY 1251 CCTCAGAAATTTGTTCTGGGTTATAGCTTCAATTCCTGTTTGGAAATTTTGTGCTAT 1310
 DB 1167 CCTCAGAAATTTGTTCTGGGTTATAGCTTCAATTCCTGTTTGGAAATTTTGTGCTAT 1226
 QY 1311 TGGCATGAGATCTTTCATTAAGCTGAATAATACACGCTATGTCATCAAAATCCT 1370
 DB 1227 TGGCATGAGATCTTTCATTAAGCTGAATAATACACGCTATGTCATCAAAATCCT 1286
 QY 1371 TTGTTGCTGATGCTGCTGATGAGGTTTAACTGTTCTTTGTTGGCAATTTGCAATTA 1430
 DB 1287 TTGTTGCTGATGCTGCTGATGAGGTTTAACTGTTCTTTGTTGGCAATTTGCAATTA 1346
 QY 1431 ATACCGAGGGCAATATACAGAAATATGCTGCTGATGATGAGAGAGCTGAGCGCCCT 1490
 DB 1347 ATACCGAGGGCAATATACAGAAATATGCTGCTGATGATGAGAGAGCTGAGCGCCCT 1406
 QY 1491 CATGGGTTCTGGGCACTGCTGTCACCGAAGTCTGTTCTGCTACTGACTGACTTATGAC 1550
 DB 1407 CATGGGTTCTGGGCACTGCTGTCACCGAAGTCTGTTCTGCTACTGACTGACTTATGAC 1466
 QY 1551 TTTGGAGAGTTCTGCTGCTATGTTCTTCCCTTCAATCAATTCGACCTGAGAAACGGCA 1610
 DB 1467 TTTGGAGAGTTCTGCTGCTATGTTCTTCCCTTCAATCAATTCGACCTGAGAAACGGCA 1526
 QY 1611 GACCTCAGTATCTGCTATTTGCAATCTGATGAGGCGGATTTTAAATAGCTGAATTCAT 1670
 DB 1527 GACCTCAGTATCTGCTATTTGCAATCTGATGAGGCGGATTTTAAATAGCTGAATTCAT 1586
 QY 1671 TTGGAATTAAGATTAATTTGGAAACCTTTTAATGGGAAAAATGAGATATGTTCCCACTTTA 1730

DB 1587 TTGGAATTAAGATTAATTTTGGAAACTTTTATGGAAAAATGAGATATGTTCCCACTTTA 1646
 QY 1721 TTATGACCAAAACAGAAATATTGGAAGCAAAAGGATTTCTTGGAAATTTTCTAGTGT 1790
 DB 1647 TTATGACCAAAACAGAAATATTGGAAGCAAAAGGATTTCTTGGAAATTTTCTAGTGT 1706
 QY 1791 GAACCTGGCTTTTTCATCATATGCTTTTCCATATTAATTAATGTTCTGTCATCA 1850
 DB 1707 GAACCTGGCTTTTTCATCATATGCTTTTCCATATTAATTAATGTTCTGTCATCA 1766
 QY 1851 AAAAACCCTTTCAGACCAACAGAAATGTAAGAAATGTTTGGAAAGAGGTCGTTGC 1910
 DB 1767 AAAAACCCTTTCAGACCAACAGAAATGTAAGAAATGTTTGGAAAGAGGTCGTTGC 1826
 QY 1911 AAATGTTTCTTTTAAATAGTGTCTCTGATGCCATCTGCTGATTTCTGAT 1970
 DB 1827 AAATGTTTCTTTTAAATAGTGTCTCTGATGCCATCTGCTGATTTCTGAT 1886
 QY 1971 TAAATTCCTTCCCTCTCCGGTGGAAATACAGACCAATGACTTCTGATAGTAT 2030
 DB 1887 TAAATTCCTTCCCTCTCCGGTGGAAATACAGACCAATGACTTCTGATAGTAT 1946
 QY 2031 TTTTTCCTTCAGTTAAAGTCTTGAATCCATCTCTATCTCTACAACTT 2090
 DB 1947 TTTTTCCTTCAGTTAAAGTCTTGAATCCATCTCTATCTCTACAACTT 2006
 QY 2091 TTTTAAAGCAAGTTGAAACAGCTGCTGACCAACATCAGAGAAATCAATTTCAAAAT 2150
 DB 2007 TTTTAAAGCAAGTTGAAACAGCTGCTGACCAACATCAGAGAAATCAATTTCAAAAT 2066
 QY 2151 TAAAAAAGGTTTATCTACATCATGTTGATGATAGAGATCTCTCTCCGTAAGT 2210
 DB 2067 TAAAAAAGGTTTATCTACATCATGTTGATGATGAGATCTCTCTCCGTAAGT 2126
 QY 2211 TGGGCTTTTGAACAAATTAACATTTGAGACAGATTAATGAACAGTTTCTAG 2265
 DB 2127 TGGGCTTTTGAACAAATTAACATTTGAGACAGATTAATGAACAGTTTCTAG 2181

RESULT 7
 ABL40191
 ID ABL40191 standard; cDNA; 1545 BP.
 XX
 AC ABL40191;
 XX
 DT 23-MAY-2002 (first entry)
 XX
 DE Human G protein-coupled receptor TGR17-2 encoding cDNA SEQ ID NO:6.
 XX
 KW Human; G protein-coupled receptor; TGR17-2; neotropic; anti-inflammatory;
 KW vasotropic; immunomodulator; cytoskeletal; gene therapy; protein therapy;
 KW neurologic; inflammatory; circulatory; degenerative; immune system;
 KW digestive disease; cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS 1..1545 Location/Qualifiers
 FT FT /*tag= a
 FT FT /partial
 FT FT /product= "TGR17-2"
 FT FT /note= "no stop codon given"
 XX
 PN MO200204640-A1.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001, 2001WO-JP05878.
 XX
 PR 07-JUL-2000, 2000JP-0211989.
 XX
 PR 18-DEC-2000, 2000JP-0383794.
 XX

(TAKE) TAKEDA CHEM IND LTD.
 Moriya T, Ito T, Shintani Y, Miyajima N;
 WPI: 2002-179706/23.
 P-PSDB: ABB06251.
 G-protein coupled receptor protein TGR17 of human origin and DNA
 encoding it for diagnosis and treatment of cancer and circulatory and
 other diseases associated with its expression -
 Claim 7; Page 118-119; 145pp; Japanese.
 The present invention describes a human guanine nucleotide binding
 protein (G protein) coupled receptor protein designated TGR17, which has
 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
 additional residues at the N-terminal. The TGR17 proteins have nootropic,
 anti-inflammatory, vasotropic, immunomodulator and cytoskeletal activities.
 The TGR17 polynucleotides and protein can be used in gene therapy and
 protein therapy. G-protein coupled receptor proteins are cell membrane
 proteins mediating the cellular response to a large variety of signalling
 molecules. The TGR17 polynucleotides and proteins can be used in the
 diagnosis, treatment and prevention of diseases including neurological,
 inflammatory, circulatory and prevention of diseases including neurological,
 diseases and cancer. The present sequence encodes human TGR17-2 from the
 present invention.
 Sequence 1545 BP; 429 A; 320 C; 290 G; 506 T; 0 other;
 Query March 68.1%; Score 1543.4; DB 24; Length 1545;
 Basic Local Similarity 99.9%; Pred. No. 0;
 Matches 1544; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 718 ATGGTATATACCTCTTGAAGCTCTTCCAGAGAGATGTGCGCAAAAGCTCACTC 777
 1 ATGGTATATACCTCTTGAAGCTCTTCCAGAGAGATGTGCGCAAAAGCTCACTC 60
 778 AACTGGGTGATTTGGAAGGCAATAGATTAAGTATCTACAAATTTCTAGCTTCG 837
 61 AACTGGGTGATTTGGAAGGCAATAGATTAAGTATCTACAAATTTCTAGCTTCG 120
 838 TGGCAATTCGCTCAAGTCTGTTCTGCTGAGAAATCAATTTGTTTTCGAGAG 897
 121 TGGCAATTCGCTCAAGTCTGTTCTGCTGAGAAATCAATTTGTTTTCGAGAG 180
 898 ACATTTCTTCAATTAATAATTTAGAGAGCTGATCTGTAGCAATAGATTAAG 957
 181 ACATTTCTTCAATTAATAATTTAGAGAGCTGATCTGTAGCAATAGATTAAG 240
 958 CTATCACTCACTTTTAAAGCTTGAAGCTTCAAAAGCTGAACCTGTCACTCAAT 1017
 241 CTATCACTCACTTTTAAAGCTTGAAGCTTCAAAAGCTGAACCTGTCACTCAAT 300
 1018 CCTCTTAATGATCTTCAAGCAAGTGAAGCTTAAACCACTTCACTCTAGAC 1077
 301 CCTCTTAATGATCTTCAAGCAAGTGAAGCTTAAACCACTTCACTCTAGAC 360
 1078 CTGAAAGATAGATTCCAATATTAACACAGATGTTTCAACCACTTAAGATCTT 1137
 361 CTGAAAGATAGATTCCAATATTAACACAGATGTTTCAACCACTTAAGATCTT 420
 1138 TCTCACTTAATTTCAAAACCTTGAATCTGTCTTCAATGCTCCCATGTCGAATGT 1197
 421 TCTCACTTAATTTCAAAACCTTGAATCTGTCTTCAATGCTCCCATGTCGAATGT 480
 1198 ATGCCCTTGAGAGAGGATTTCTTCAATTAAGACCTTGGCTTAACAATCTCTAGA 1257
 481 ATGCCCTTGAGAGAGGATTTCTTCAATTAAGACCTTGGCTTAACAATCTCTAGA 540
 1258 ATATTTGTCTGGGTTATAGCTTCTTCACTTCACTGCTTTGGAATCTTTTGTGATGGCAGT 1317
 541 ATATTTGTCTGGGTTATAGCTTCTTCACTTCACTGCTTTGGAATCTTTTGTGATGGCAGT 600

1318 AGATCTTCAATTAAGCTGAATAATACAACTACGCTATGTCATCAAAATCTTTGTGT 1377
 601 AGATCTTCAATTAAGCTGAATAATACAACTACGCTATGTCATCAAAATCTTTGTGT 660
 1378 GCTGATTCCTGATGGGTGTTTACTTGTCTTTGTTGGCAATTTTGCATTAATAATACCGA 1437
 661 GCTGATTCCTGATGGGTGTTTACTTGTCTTTGTTGGCAATTTTGCATTAATAATAACCGA 720
 1438 GGGCAGTATCAGAGATATGCTGCTGTGTGATGAGAGCGTGCAGTCCGCTCATGGGG 1497
 721 GGGCAGTATCAGAGATATGCTGCTGTGTGATGAGAGCGTGCAGTCCGCTCATGGGG 780
 1498 TTCTGCGCATGCTGTCCACGAGAGTCTGTCTGTCTGCTACTGACTTGAATTTGAG 1557
 781 TTCTGCGCATGCTGTCCACGAGAGTCTGTCTGTCTGCTACTGACTTGAATTTGAG 840
 1558 AAGTCTGCTGATGCTGCTTCCCTTGTAGTAACTTGAACCTGGAATAACGCGACCTCA 1617
 841 AAGTCTGCTGATGCTGCTTCCCTTGTAGTAACTTGAACCTGGAATAACGCGACCTCA 900
 1618 GTCACTCTCATTTGCAATCTGATGAGAGGATTTTAACTGTAAATTTCAATTTGAG 1677
 901 GTCACTCTCATTTGCAATCTGATGAGAGGATTTTAACTGTAAATTTCAATTTGAG 960
 1678 AAGGATTAATTTGGAACCTTTATGGAATAATGAGATATGTTCCACTTAATATGAC 1737
 961 AAGGATTAATTTGGAACCTTTATGGAATAATGAGATATGTTCCACTTAATATGAC 1020
 1738 CAAACAGAAATATTTGGAACCAAGGATTTCTTGTGAATTTTCTAGGTGTAACCTTG 1797
 1021 CAAACAGAAATATTTGGAACCAAGGATTTCTTGTGAATTTTCTAGGTGTAACCTTG 1080
 1798 CTGGCTTTTCTCATCATTTGTTTCTTCTATATTAATTAATGTTCTGTTTCAATCAAAAACC 1857
 1081 CTGGCTTTTCTCATCATTTGTTTCTTCTATATTAATTAATGTTCTGTTTCAATCAAAAACC 1140
 1858 GCCTTGCAACAACAAGATTAAGAAATTTTGAAGAGAGGTGCTGTTCGAATCGT 1917
 1141 GCCTTGCAACAACAAGATTAAGAAATTTTGAAGAGAGGTGCTGTTCGAATCGT 1200
 1918 TTCTTTTATTAAGTGTCTGTGATGCAATCTGCTGATTTCTGTATTTGATTAATATC 1977
 1201 TTCTTTTATTAAGTGTCTGTGATGCAATCTGCTGATTTCTGTATTTGATTAATATC 1260
 1978 CTTTCCCTCTCCGGGGTGAATATACGACACATGACTTCTGATATGATTTTTC 2037
 1261 CTTTCCCTCTCCGGGGTGAATATACGACACATGACTTCTGATATGATTTTTC 1320
 2038 CTTCAGTAAACAAGTCTTGAATCCAAATCTCTAATCTCTCAACCAACTTTTAAAG 2097
 1321 CTTCAGTAAACAAGTCTTGAATCCAAATCTCTAATCTCTCAACCAACTTTTAAAG 1380
 2098 GACAAAGTGAACAAGCTGCTGCAACAATCAAGAGAAATCAATTTTCAAAATTAATAA 2157
 1381 GACAAAGTGAACAAGCTGCTGCAACAATCAAGAGAAATCAATTTTCAAAATTAATAA 1440
 2158 AAAAGTTATCTATCATCATTTGTGTGATAGAGACTCTTCCCTGAAATCTGGGTT 2217
 1441 AAAAGTTATCTATCATCATTTGTGTGATAGAGACTCTTCCCTGAAATCTGGGTT 1500
 2218 TTGAACAAATTAACACTTGGAGACAGTATATGAACCAAGTTTC 2262
 1501 TTGAACAAATTAACACTTGGAGACAGTATATGAACCAAGTTTC 1545
 RESULT 8
 ABL40192
 ID ABL40192 standard; cDNA; 1473 BP.
 XX ABL40192:
 AC
 XX
 XX
 DT 23-MAY-2002 (first entry)
 XX

Oy		2038	CTTCAGTTACAGAGCGTTTGATCAATCCGTAATACTGCACAACCACTTTTAAg	2097
Db		1249	CTTCAGTTAACAGTGTCTTTGAATCCATTCTCTAATCTCACAACCACCTTTTTAAg	1308
Oy		2098	GACAAGTTGAAAACAGCTGCTGCACAAAATCATGAGGAAATCAATTTTCAAATTAATAAAA	2157
Db		1309	GACAAGTTGAAAACAGCTGCTGCACAAAATCATGAGGAAATCAATTTTCAAATTAATAAAA	1368
Oy		2158	AAAAGTTATATCTACATCCATGTGTGGATAGAAGACTCTCTTCCCTGAAAATTGGGGTT	2217
Db		1369	AAAAGTTATATCTACATCCATGTGTGGATAGAAGACTCTCTTCCCTGAAAATTGGGGTT	1428
Oy		2218	TGGAACAATAATPAACCTTGAGACAGATTAATGAAAACAGTTTCC	2262
Db		1429	TTGAACAATAATPAACCTTGAGACAGATTAATGAAAACAGTTTCC	1473
RESULT 9				
ABL40188	ID	ABL40188	standard; cDNA; 1830 BP.	
XX	AC	ABL40188;		
DT	DT	23-MAY-2002	(first entry)	
XX	XX			
XX	DE	Human G protein-coupled receptor TGR17-1 encoding cDNA SEQ ID NO:2.		
KW	KW	Human; G protein-coupled receptor; TGR17-1; nootropic; antiinflammatory;		
KM	KM	vasotrophic; immunomodulator; cytostatic; gene therapy; protein therapy;		
KX	KX	neurological; inflammatory; circulatory; degenerative; immune system;		
XX	KX	digestive disease; cancer; gene; ss.		
OS	OS	Homo sapiens.		
XX	XX			
FH	Key	Location/Qualifiers		
FT	CDS	1..1830		
FT		/tag= a		
FT		/partial		
FT		/product= "TGR17-1"		
PN		/note= "no stop codon given"		
XX	XX			
XX	XX	WO200204640-AI.		
PB	PB	17-JAN-2002.		
XX	XX			
PF	PF	06-JUN-2001; 2001WO-JP05878.		
XX	XX			
PR	PR	07-JUL-2000; 2000JP-0211989.		
XX	XX			
PR	PR	18-DEC-2000; 2000JP-0383794.		
XX	XX			
PA	PA	(TAKE) TAKEDA CHEM IND LTD.		
XX	XX			
PI	PI	Moriya T, Ito T, Shintani Y, Miyajima N;		
DR	DR	WPI; 2002-179706/23.		
XX	XX			
PT	PT	P-PsDB; ABB06250.		
XX	XX			
PT	PT	G-protein coupled receptor protein TGR17 of human origin and DNA		
XX	XX	encoding it for diagnosis and treatment of cancer and circulatory and		
XX	XX	other diseases associated with its expression -		
PS	PS	Claim 7; Page 114-115; 145pp; Japanese.		
XX	XX			
CC	CC	The present invention describes a human guanine nucleotide binding		
CC	CC	protein (G-protein) coupled receptor protein designated TGR17, which has		
CC	CC	5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having		
CC	CC	additional residues at the N-terminal. The TGR17 proteins have nootropic,		
CC	CC	antiinflammatory, vasotrophic, immunomodulator and cyostatic activities.		
CC	CC	The TGR17 polynucleotides and protein can be used in gene therapy and		
CC	CC	protein therapy. G-protein coupled receptor proteins are cell membrane		
CC	CC	proteins mediating the cellular response to a large variety of signalling		
CC	CC	molecules. The TGR17 polynucleotides and proteins can be used in the		
CC	CC	diagnosis, treatment and prevention of diseases including neurological,		

CC	inflammatory, circulatory, degenerative, immune system and digestive	
CC	diseases and cancer. The present sequence encodes human TGR17-1 from the	
CC	present invention.	
XX		
Sequence	1830 BP; 516 A; 377 C; 351 G; 586 T; 0 other;	
Query Match	61.1%; Score 1384.8; DB 24; Length 1830;	
Best Local Similarity	80.8%; Pred. No. 0;	
Matches 1828; Conservative	0; Mismatches 2; Indels 432; Gaps 1	
QY	1 ATGATTTGTTTTCTGGTTTTTAAACATCTCTTCAGCCTCAGATTGATTACATGTTCTTT	60
DB	1 ATGATTTGTTTTCTGGTTTTTAAACATCTCTTCAGCCTCAGATTGATTACATGTTCTTT	60
QY	61 CTACTTCATTCATCGTTCGTATCATCAATGTCAAAAGATTTTGCATGCACTCAAGSTAGCATG	120
DB	61 CTACTTCATTCATCGTTCGTATCATCAATGTCAAAAGATTTTGCATGCACTCAAGSTAGCATG	120
QY	121 ATCACTCCCTCATGCGCAAAAAGATATTTTCCCTGTGGAAATCTTACAGAGTCTTACCC	180
DB	121 ATCACTCCCTCATGCGCAAAAAGATATTTTCCCTGTGGAAATCTTACCAAGTCTTACCC	180
QY	181 CGAGCTTTTCACTGTGATGCGCAAGATGACATGTGGGAACGGGGCCGAGGAAGAATCT	240
DB	181 CGAGCTTTTCACTGTGATGCGCAAGATGACATGTGGGAACGGGGCCGAGGAAGAATCT	240
QY	241 GGTGACATTAATGATGATGGCGCAACATATTTGGCAGATGCAATGAAATGCTAACAGCGTG	300
DB	241 GGTGACATTAATGATGATGGCGCAACATATTTGGCAGATGCAATGAAATGCTAACAGCGTG	300
QY	301 GCCTTAAACAAGAGTGCCTTTCTTAAACAAGTATCCAAATGCTGTGACTGCAGAAAGAACT	360
DB	301 GCCTTAAACAAGAGTGCCTTTCTTAAACAAGTATCCAAATGCTGTGACTGCAGAAAGAACT	360
QY	361 GAATTTGAATGTGTAAATGATGACTTAAAGTGTGCGGATGATTCTTAAAGATGAGA	420
DB	361 GAATTTGAATGTGTAAATGATGACTTAAAGTGTGCGGATGATTCTTAAAGATGAGA	420
QY	421 TTACTGTCTCTTAAGAAAAACAAATTCACAGTCTTCCAGATTAAGTTTTCATCAATAC	480
DB	421 TTACTGTCTCTTAAGAAAAACAAATTCACAGTCTTCCAGATTAAGTTTTCATCAATAC	480
QY	481 ACAAACCTTAAAAAATATTTCTTACAGATATGTGCTTATGACACATATCCAGAAAGCA	540
DB	481 ACAAACCTTAAAAAATATTTCTTACAGATATGTGCTTATGACACATATCCAGAAAGCA	540
QY	541 TTTTGTGATTATGTAATCTGCAATATATATATCTCAACCAACAGTGCATCAACCTTC	600
DB	541 TTTTGTGATTATGTAATCTGCAATATATATATCTCAACCAACAGTGCATCAACCTTC	600
QY	601 AGACCTGGAATATTCAAGAACTTACATCAGCTAATCTGCTTAATCTAGATGCAATCCA	660
DB	601 AGACCTGGAATATTCAAGAACTTACATCAGCTAATCTGCTTAATCTAGATGCAATCCA	660
QY	497 -----	496
QY	661 ATTAACCAAGATTTCACAGCGCTGTGTTACGGGATTAATTCCTTGTTTTCTGTCTATG	720
DB	497 -----	496
QY	721 GTTAATACTACTTAGAAGCTCTTCCCAAGCAGATGTGTGCCCAATGCTCTCAACTCAAC	780
DB	497 -----	496
QY	781 TGGGTGATTTGGAAAGCAATAGAAATTAAGTATCTCACAAATTTACGTTTCTGTGCTGC	840
DB	497 -----	496
QY	841 GATTGCTACAGTGTCTTTCTGCTAGAAATCAAAATGGTTTTGTTCCAGAGAAAGA	900
DB	497 -----	496
QY	901 TTTTCTTCAATTAATAATTTAGAGAACTGATCTGTCTAGCAATACGATTAACGAGACTA	960
DB	497 -----	496
QY	961 TTTTCTTCAATTAATAATTTAGAGAACTGATCTGTCTAGCAATACGATTAACGAGACTA	960
DB	497 -----	496
QY	997 -----	996
DB	997 -----	996

```
QY 961 TCACCTCACTTTTAAAGCTTGAAGCTTCTACAAAAGCTGAACCTGTCAATCCT 1020
DB 529 TCACCTCACTTTTAAAGCTTGAAGCTTCTACAAAAGCTGAACCTGTCAATCCT 588
QY 1021 CTTATGATCTTCAACAAGACAGTTTGAAGCTTAAACAATTCAAGCTCTAGACCTG 1080
DB 589 CTTATGATCTTCAACAAGACAGTTTGAAGCTTAAACAATTCAAGCTCTAGACCTG 648
QY 1081 GAAAGATAGAGATTCACAAATATAAACACAGATGTTTCAACCCATGAAGATCTTCT 1140
DB 649 GAAAGATAGAGATTCACAAATATAAACACAGATGTTTCAACCCATGAAGATCTTCT 708
QY 1141 CACATTTATTTCAAAAACCTTGTGATCTGTCTCTATGCTCCCATGTCCGATATGTATG 1200
DB 709 CACATTTATTTCAAAAACCTTGTGATCTGTCTCTATGCTCCCATGTCCGATATGTATG 768
QY 1201 CCCTTGACGAGCGGACATTTCTTCAATTGAGACCTCTTGAGCTACAAATCTCTCAGAAATA 1260
DB 769 CCCTTGACGAGCGGACATTTCTTCAATTGAGACCTCTTGAGCTACAAATCTCTCAGAAATA 828
QY 1261 TTTGTCTGGGTTATAGCTTTCATTTACCTGCTTGAAGATCTTTTGTCAATTGGCATGAGA 1320
DB 829 TTTGTCTGGGTTATAGCTTTCATTTACCTGCTTGAAGATCTTTTGTCAATTGGCATGAGA 888
QY 1321 TCTTTTATTAAGCTGAAAAATACACTCAGCTATGTCCATCAAAATCTTTGTGTGCT 1380
DB 889 TCTTTTATTAAGCTGAAAAATACACTCAGCTATGTCCATCAAAATCTTTGTGTGCT 948
QY 1381 GATTCGCTGATGGGCTGTTTACTGTTCTTGTGGAATTTTGCATTTAAATATCCAGAGG 1440
DB 949 GATTCGCTGATGGGCTGTTTACTGTTCTTGTGGAATTTTGCATTTAAATATCCAGAGG 1008
QY 1441 CAGTATCAGAAATAGTCTTGTCTGTGATGAGAGAGCTGAGTCCGCTCATGGGCTTC 1500
DB 1009 CAGTATCAGAAATAGTCTTGTCTGTGATGAGAGAGCTGAGTCCGCTCATGGGCTTC 1068
QY 1501 CTGGCCATGCTGTCACACCGAAGTCTGCTGCTGCTCTGACCTGACTTGAATTTGGAGAG 1560
DB 1069 CTGGCCATGCTGTCACACCGAAGTCTGCTGCTGCTCTGACCTGACTTGAATTTGGAGAG 1128
QY 1561 TTCCTGATCATTTGCTTCCCTTCAATACATTCGACCTTGAAAAAGCGAGACCTCAGTC 1620
DB 1129 TTCCTGATCATTTGCTTCCCTTCAATACATTCGACCTTGAAAAAGCGAGACCTCAGTC 1188
QY 1621 ATCTCATTTGCAATCTGATGAGCGGATTTTAAATAGCTGTAAATCTTTTGAATTAAG 1680
DB 1189 ATCTCATTTGCAATCTGATGAGCGGATTTTAAATAGCTGTAAATCTTTTGAATTAAG 1248
QY 1681 GATTATTTTGGAAACCTTTATGGAAGAAAAATGGAATGTTTCCCACTTTATATGACCAA 1740
DB 1249 GATTATTTTGGAAACCTTTATGGAAGAAAAATGGAATGTTTCCCACTTTATATGACCAA 1308
QY 1741 ACAGAAGATATTGGAACCAAGGATATTCTTGGAAATTTTCTAGGTGGAATCTGCTG 1800
DB 1309 ACAGAAGATATTGGAACCAAGGATATTCTTGGAAATTTTCTAGGTGGAATCTGCTG 1368
QY 1801 GCTTTTCTCATCTGTTGTTTCTGTATATTAATGTTCTGTTCCATTTCAAAAAAGCGCC 1860
DB 1369 GCTTTTCTCATCTGTTGTTTCTGTATATTAATGTTCTGTTCCATTTCAAAAAAGCGCC 1428
QY 1861 TTGCAGACCAAGAAATGTTTGGAAAGAGAGGCTGTTGCAAAATCGTTTC 1920
DB 1429 TTGCAGACCAAGAAATGTTTGGAAAGAGAGGCTGTTGCAAAATCGTTTC 1488
QY 1921 TTTTATTAAGTGTCTGTATGTCATCTGATCTGTGATTTGTAGTTAAATCTT 1980
DB 1489 TTTTATTAAGTGTCTGTATGTCATCTGATCTGTGATTTGTAGTTAAATCTT 1548
QY 1981 TCCCTCTTCGGGGTGAATAACAGACAAATGCTCTGAGATGATTTTTCCT 2040
DB 1549 TCCCTCTTCGGGGTGAATAACAGACAAATGCTCTGAGATGATTTTTCCT 1608
```

```
QY 2041 CCAGTTAACAGTGTCTTGAATCCATCTCTATTAATCTCAGAACCACTTTTAAAGAC 2100
DB 1609 CCAGTTAACAGTGTCTTGAATCCATCTCTATTAATCTCAGAACCACTTTTAAAGAC 1668
QY 2101 AAGTTGAACAGCTGCTCAGAACCAATCAGAGGAATCAATTTTCAAAATTTAAAAAAA 2160
DB 1669 AAGTTGAACAGCTGCTCAGAACCAATCAGAGGAATCAATTTTCAAAATTTAAAAAAA 1728
QY 2161 AGTTATATCATCATCTTGTGTGATGATGAGACTCTCTTCCCTGAAAATCTGGGCTTTG 2220
DB 1729 AGTTATATCATCATCTTGTGTGATGATGAGACTCTCTTCCCTGAAAATCTGGGCTTTG 1788
QY 2221 AACCAATATACACTTGAGACAGTATTAATGAACAGTTTCC 2262
DB 1789 AACCAATATACACTTGAGACAGTATTAATGAACAGTTTCC 1830

RESULT 10
ABT33349
ID ABT33349 standard; DNA; 1121 BP.
XX
AC ABT33349;
XX
DT 22-MAY-2003 (first entry)
XX
DE NOVX DNA sequence SEQ ID NO 13.
XX
XX Hepatotropic; immunosuppressive; cardiant; hypertensive; tranquilizer;
XX vulnerey; viricide; antibacterial; protozoicide; fungicide; nootropic;
XX antiparasitic; neuroprotective; cerebroprotective; antiparkinsonian;
XX anticonvulsant; antidiabetic; analgesic; dermatological; keratolytic;
XX antiseborrheic; antihemetic; antiarthritic; antiinflammatory; anti-HIV;
XX cyostatic; antiaesthetic; antiporiatic; hypotensive; osteopathic;
XX antulcer; anorectic; antidiabetic; antiallergic; haemostatic;
XX neuroleptic; antidepressant; antifertility; NOVX; human disease;
XX NOVX-associated disorder; trauma; viral; bacterial; fungal; protozoal;
XX parasitic infection; Alzheimer's disease; stroke; forensic biology;
XX immunogen; non-human transgenic animal; gene therapy; gene; de.
XX
OS Unidentified.
XX
PN WO200281517-A2.
XX
PD 17-OCT-2002.
XX
PF 22-JAN-2002; 2002WO-US02064.
XX
XX 19-JAN-2001; 2001US-262892P.
PR 23-JAN-2001; 2001US-263598P.
PR 24-JAN-2001; 2001US-263799P.
PR 25-JAN-2001; 2001US-264117P.
PR 25-JAN-2001; 2001US-264139P.
PR 26-JAN-2001; 2001US-264478P.
PR 30-JAN-2001; 2001US-263351P.
PR 02-MAR-2001; 2001US-272870P.
PR 14-MAR-2001; 2001US-275927P.
PR 14-MAR-2001; 2001US-275980P.
PR 15-MAR-2001; 2001US-276449P.
PR 20-MAR-2001; 2001US-277358P.
PR 23-MAR-2001; 2001US-278151P.
PR 29-MAR-2001; 2001US-279857P.
PR 20-APR-2001; 2001US-285140P.
PR 20-APR-2001; 2001US-285141P.
PR 30-APR-2001; 2001US-287484P.
PR 17-MAY-2001; 2001US-291701P.
PR 08-JUN-2001; 2001US-296606P.
PR 10-JUL-2001; 2001US-304353P.
PR 10-JUL-2001; 2001US-304355P.
PR 12-JUL-2001; 2001US-304886P.
PR 09-AUG-2001; 2001US-311289P.
PR 13-AUG-2001; 2001US-311975P.
PR 16-AUG-2001; 2001US-312937P.
PR 18-OCT-2001; 2001US-330227P.
```


ID AAS07943 standard; cDNA, 1068 BP.
AC AAS07943;
XX 23-OCT-2001 (first entry)
XX Human cDNA encoding G-protein coupled receptor, hrup16.
DE Human; G-protein coupled receptor; GPCR; hrup16; agonist;
KW Inverse agonist; lung cancer; ss.
XX
OS Homo sapiens.
PH Key Location/Qualifiers
FT CDS 1..1068
FT /*tag= a
FT /product= "hrup16"
XX
PN MO200136471-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000MO-US31509.
XX
PR 17-NOV-1999; 99US-0166088.
PR 17-NOV-1999; 99US-0166099.
PR 17-NOV-1999; 99US-0166369.
PR 23-DEC-1999; 99US-0171900.
PR 23-DEC-1999; 99US-0171901.
PR 23-DEC-1999; 99US-0171902.
PR 11-FEB-2000; 2000US-0181749.
PR 14-MAR-2000; 2000US-0189258.
PR 14-MAR-2000; 2000US-0189259.
PR 10-APR-2000; 2000US-0195898.
PR 10-APR-2000; 2000US-0195899.
PR 10-APR-2000; 2000US-0196078.
PR 28-APR-2000; 2000US-0200419.
PR 12-MAY-2000; 2000US-0203630.
PR 12-JUN-2000; 2000US-0210741.
PR 12-JUN-2000; 2000US-0210982.
PR 21-AUG-2000; 2000US-0226760.
PR 26-SEP-2000; 2000US-0235418.
PR 26-SEP-2000; 2000US-0235779.
PR 20-OCT-2000; 2000US-0242332.
PR 20-OCT-2000; 2000US-0242343.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Dang HT, Lowitz KP;
XX WPI; 2001-355616/37.
DR P-PSDB; AAU04370.
XX
PT Endogenous and non-endogenous versions of human G-protein coupled
PT receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -
XX
PS Claim 35; Page 104-105; 159pp; English.
XX
XX The sequence encodes a human G-protein coupled receptor (GPCR),
CC hup16. The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilized to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.
XX
SQ Sequence 1068 BP; 278 A; 215 C; 214 G; 361 T; 0 other;
Query Match 47.1%; Score 1066.4; DB 22; Length 1068;

Best Local Similarity 99.9%; Pred. No. 4,9e-275;
Matches 1067; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1198 ATGCCCTTGACGAGCGCATTTCTTCATTGAGACCTTGGCTAACAAATTCCTCAGA 1257
DB 1 ATGCCCTTGACGAGCGCATTTCTTCATTGAGACCTTGGCTAACAAATTCCTCAGA 60
QY 1258 ATATTGTCTGGATTATAGCTTTCATTACCTGCTTGGAAATCTTTTGTCAATGGCAG 1317
DB 61 ATATTGTCTGGATTATAGCTTTCATTACCTGCTTGGAAATCTTTTGTCAATGGCAG 120
QY 1318 AGATCTTTCATTAAAGCTGAAATCAACTCAGCTATGCTCATCAAAATCTTGTGTCT 1377
DB 121 AGATCTTTCATTAAAGCTGAAATCAACTCAGCTATGCTCATCAAAATCTTGTGTCT 180
QY 1378 GCTGATGCTGATGAGGTGTTTACTGTTCTTTGTTGGCATTTGATTAATAATACGA 1437
DB 181 GCTGATGCTGATGAGGTGTTTACTGTTCTTTGTTGGCATTTGATTAATAATACGA 240
QY 1438 GGGCAGTATCAGAAATATGCTTGTCTGTGAGATGAGAGCGTGCAGTCCGCTCATGGG 1497
DB 241 GGGCAGTATCAGAAATATGCTTGTCTGTGAGATGAGAGCGTGCAGTCCGCTCATGGG 300
QY 1498 TTCTGGGCAATGCTGTCCACGGAAGCTCTGTTGCTACTGACCTACTTGACTTTGAG 1557
DB 301 TTCTGGGCAATGCTGTCCACGGAAGCTCTGTTGCTACTGACCTACTTGACTTTGAG 360
QY 1558 AAGTTCCTGTCATGTCCTTCCCTTCAGTAACTTGCACCTGGAAGCGGACCTCA 1617
DB 361 AAGTTCCTGTCATGTCCTTCCCTTCAGTAACTTGCACCTGGAAGCGGACCTCA 420
QY 1618 GTTCATCTCATTTGCATCTGATGCGGGATTTTATATGCTGATATTCATTTGGAT 1677
DB 421 GTTCATCTCATTTGCATCTGATGCGGGATTTTATATGCTGATATTCATTTGGAT 480
QY 1678 AAGATTAATTTTGGAACTTTATGAGAAATAGAGATGTTTCCACTTATATATGAC 1737
DB 481 AAGATTAATTTTGGAACTTTATGAGAAATAGAGATGTTTCCACTTATATATGAC 540
QY 1738 CAACAGAAATATTTGAGAAAGGATATCTTTGAAATTTTCTAGGTGAACTTG 1797
DB 541 CAACAGAAATATTTGAGAAAGGATATCTTTGAAATTTTCTAGGTGAACTTG 600
QY 1798 CTGGCTTCTCATCATCTGATGTTTCTTATATGATGCTGCTGCTCAATCAAAAC 1857
DB 601 CTGGCTTCTCATCATCTGATGTTTCTTATATGATGCTGCTGCTCAATCAAAAC 660
QY 1858 GCCTTGACAGCAAGAAATGAAATGATTTGTTGAAAGAGTGGCTGTGCAATCGT 1917
DB 661 GCCTTGACAGCAAGAAATGAAATGATTTGTTGAAAGAGTGGCTGTGCAATCGT 720
QY 1918 TTCTTTTATATAGTGTCTCTGATGCTGCTGATCTGATATTTAGTTAAATC 1977
DB 721 TTCTTTTATATAGTGTCTCTGATGCTGCTGATCTGATATTTAGTTAAATC 780
QY 1978 CTTCCTCTCTCCGGTGAATATACAGACAAATGACTTCCGTGATAGATTTTTC 2037
DB 781 CTTCCTCTCTCCGGTGAATATACAGACAAATGACTTCCGTGATAGATTTTTC 840
QY 2038 CTTCAGTTAAGATGCTTGAATCAATCTCTATACTCTGACAAACCACTTTTAAAG 2097
DB 841 CTTCAGTTAAGATGCTTGAATCAATCTCTATACTCTGACAAACCACTTTTAAAG 900
QY 2098 GACAAATTGAAACAGCTGCTGCAAAACATCAGAGAAATCAATTTTCAAAATTA 2157
DB 901 GACAAATTGAAACAGCTGCTGCAAAACATCAGAGAAATCAATTTTCAAAATTA 960
QY 2158 AAAAGTTATCTACATCAATGATGAGATGAGAGACTCTCTCCCTGAAACTTGGGGTT 2217
DB 961 AAAAGTTATCTACATCAATGATGAGATGAGAGACTCTCTCCCTGAAACTTGGGGTT 1020
QY 2218 TTGAACAAATTAACCTTGAGACAGTATATGAACCAATTTCTTCTAG 2265

Db 1021 TTGAACAAATTAACCTTGAGACGATATATGAACCACTTCTAG 1068

RESULT 12

ID ABL40194 standard; cDNA; 1065 BP.

XX ABL40194;

XX 23-MAY-2002 (first entry)

XX Human G protein-coupled receptor TGR17-4 encoding cDNA SEQ ID NO:11.

XX Human; G protein-coupled receptor; TGR17-4; nootropic; antiinflammatory;

KM vasoactive; immunomodulator; cytosolic; gene therapy; protein therapy;

KM neurological; inflammatory; circulatory; degenerative; immune system;

KM digestive disease; cancer; gene; ss.

XX Homo sapiens.

XX Key location/Qualifiers

XX CDS 1.1065

XX /tag= a

XX /partial

XX /product= "TGR17-4"

XX /note= "no stop codon given"

XX MO200204640-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-JP05878.

XX 07-JUL-2000; 2000JP-0211989.

XX 18-DEC-2000; 2000JP-0383794.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Moriya T, Ito T, Shintani Y, Miyajima N;

XX WPI; 2002-179706/23.

XX P-PSDB; ABB06253.

XX G-protein coupled receptor protein TGR17 of human origin and DNA

XX encoding it for diagnosis and treatment of cancer and circulatory and

XX other diseases associated with its expression -

XX Claim 7; Page 125-126; 145pp; Japanese.

XX The present invention describes a human guanine nucleotide binding

XX protein (G protein) coupled receptor protein designated TGR17, which has

XX 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having

XX additional residues at the N-terminal. The TGR17 proteins have nootropic,

XX antiinflammatory, vasoactive, immunomodulator and cytosolic activities.

XX The TGR17 polynucleotides and protein can be used in gene therapy and

XX protein therapy. G-protein coupled receptor proteins are cell membrane

XX molecules mediating the cellular response to a large variety of signalling

XX molecules. The TGR17 polynucleotides and proteins can be used in the

XX diagnosis, treatment and prevention of diseases including neurological,

XX inflammatory, circulatory, degenerative, immune system and digestive

XX diseases and cancer. The present sequence encodes human TGR17-4 from the

XX present invention.

XX Sequence 1065 BP; 277 A; 214 C; 213 G; 361 T; 0 other;

XX Query Match 47.0%; Score 1065; DB 24; Length 1065;

XX Best Local Similarity 100.0%; Pred. No. 1.2e-274;

XX Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1198 ATGCCCTTGACGACGACGATTTCTTCAATTTGAGACCTTGGCTAACAAATATCTCTGAGA 1257

DB 1 ATGCCCTTGACGACGACGATTTCTTCAATTTGAGACCTTGGCTAACAAATATCTCTGAGA 60

QY 1258 ATATTGCTGGGTATAGCTTTCAATTAACCTGCTTGGAAATCTTTTGTCAATGGCAGT 1317

DB 61 ATATTGCTGGGTATAGCTTTCAATTAACCTGCTTGGAAATCTTTTGTCAATGGCAGT 120

QY 1318 AGATCTTTCATTAAGCTGAAAATACACCTACGCTATGTCATCAAAAATCTTTGTGT 1377

DB 121 AGATCTTTCATTAAGCTGAAAATACACCTACGCTATGTCATCAAAAATCTTTGTGT 180

QY 1378 GCTGATTCGCTGATAGGCTGTTTACTGTTTGTGGCAATTTTGCATATTAATAACCGA 1437

DB 181 GCTGATTCGCTGATAGGCTGTTTACTGTTTGTGGCAATTTTGCATATTAATAACCGA 240

QY 1438 GGGCAGATACAGAAATAGCTGCTGCTGATGAGAGCGCTGACGCTGATGAGG 1497

DB 241 GGGCAGATACAGAAATAGCTGCTGCTGATGAGAGCGCTGACGCTGATGAGG 300

QY 1498 TTCTGGCAGATGCTGTCACCGAAAGTCTGTTCTGCTACTGACCTTGAATTTGAG 1557

DB 301 TTCTGGCAGATGCTGTCACCGAAAGTCTGTTCTGCTACTGACCTTGAATTTGAG 360

QY 1558 AAGTTCCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617

DB 361 AAGTTCCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 1618 GTCATCCTCATTTGATCTGAGATGCGGGAATTTTAACTGTAATTCATTTTGGAAAT 1677

DB 421 GTCATCCTCATTTGATCTGAGATGCGGGAATTTTAACTGTAATTCATTTTGGAAAT 480

QY 1678 AAGATTAATTTTGGAAATCTTTTATGAGAAATAGAGATGTTTCCCATTTATATGAC 1737

DB 481 AAGATTAATTTTGGAAATCTTTTATGAGAAATAGAGATGTTTCCCATTTATATGAC 540

QY 1738 CAACAGAGATATTTGAGAAAGGATTTCTTGGAAATTTTCTAGGTTGACCTG 1797

DB 541 CAACAGAGATATTTGAGAAAGGATTTCTTGGAAATTTTCTAGGTTGACCTG 600

QY 1798 CTGCTTTTCTCATCATTTGTTTCTTATATTAATTAATGTTCTGCTCATTTCAAAAAACC 1857

DB 601 CTGCTTTTCTCATCATTTGTTTCTTATATTAATTAATGTTCTGCTCATTTCAAAAAACC 660

QY 1858 GCTTGCAGACACACAGAAATGTTTGGAAAGAGTGGCTGTTGCAATTCGT 1917

DB 661 GCTTGCAGACACACAGAAATGTTTGGAAAGAGTGGCTGTTGCAATTCGT 720

QY 1918 TTTCTTTTATAGTCTGCTGATGCAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1977

DB 721 TTTCTTTTATAGTCTGCTGATGCAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 780

QY 1978 CTTTCCCTCTTCCGGTGAATATACAGACACATGACTTCTGATATGATTTTTC 2037

DB 781 CTTTCCCTCTTCCGGTGAATATACAGACACATGACTTCTGATATGATTTTTC 840

QY 2038 CTTTCAATTAACATGCTTTGATTCATCTCTTATCTGACACCACTTTTAAAG 2097

DB 841 CTTTCAATTAACATGCTTTGATTCATCTCTTATCTGACACCACTTTTAAAG 900

QY 2098 GACAAGTTGAAACGCTGCGACAAACATCAGAGAAATCAATTTCAAAAATTAATAA 2157

DB 901 GACAAGTTGAAACGCTGCGACAAACATCAGAGAAATCAATTTCAAAAATTAATAA 960

QY 2158 AAAAGTTATATCATCATTTGTTGATAGAGATCTTCTTCCGTGAAATCTGGGTT 2217

DB 961 AAAAGTTATATCATCATTTGTTGATAGAGATCTTCTTCCGTGAAATCTGGGTT 1020

QY 2218 TTGAACAAATTAACCTTGAGACGATATATGAACCACTTTC 2262

DB 1021 TTGAACAAATTAACCTTGAGACGATATATGAACCACTTTC 1065

RESULT 13

ABZ42570

ID ABZ42570 standard; DNA; 2274 BP.

XX

AC AB242570;
 XX 04-MAR-2003 (first entry)
 XX Human G protein-coupled receptor LGR7 nucleotide seq ID NO:622.
 DE
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 XX G protein-coupled receptor; modulator; antibody; immune-related disease;
 XX growth-related disease; cell regeneration-related disease; AIDS; cancer;
 XX immunological-related cell proliferative disease; autoimmune disease;
 XX Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 XX osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 XX graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 XX psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 XX mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 XX hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 XX ulcer; gene; ds.
 XX Homo sapiens.
 OS
 XX WO200261087-A2.
 PN
 XX 08-AUG-2002.
 PD
 XX 19-DEC-2001; 2001WO-US50107.
 PP
 XX 19-DEC-2000; 2000US-257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR
 XX P-PSDB; ABP81724.
 XX
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating
 PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
 PT cancer or autoimmune diseases -
 PT
 XX
 XX Disclosure; Fig 1; 523pp; English.
 PS
 XX The present invention describes antigenic peptides (I) comprising:
 CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular
 CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity
 CC or avidity for a particular GPCR. (1) can be used as GPCR modulators and
 CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
 CC an antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, memory
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention.
 XX
 SO Sequence 2274 BP; 669 A; 460 C; 421 G; 724 T; 0 other;
 Query Match 32.5%; Score 735; DB 25; Length 2274;
 Best Local Similarity 60.9%; Pred. No. 3,96-186;
 Matches 1263; Conservative 0; Mismatches 785; Indels 27; Gaps 3;

Qy	142	GGATATTTTCCCTGTGGGAATCTTACCAAGTCTTACCCGAGCTTCACTGTGATGCG	201
Db	88	GGCTATTTCCCTGTGGGAATCTTACCAAGTCTTACCCGAGCTTCACTGTGATGCG	147
Qy	202	AAGATGACGTGGGAACGGGCGGACGAGAGAAGACTGTGTGACACTAGTGTGATGCGG	261
Db	148	GTGAGACGCTCGGGAAATCAGGCCGATGAGACAACTGTGTGAGACAAATGATGTGCTC	207
Qy	262	ACCATTTTGGCAAGTGCATGGAATGCTAAC-----GCGTG	300
Db	208	ATGCAATTTGCAAAATTTTGGCAGTACTACAAATATGATCTCCATATCTTTTGTAG	267
Qy	301	GCTTAAACAGAGAGTCTTTCTAAACAGATATCCAAATGCTGTGACGTGAGAAAGACT	360
Db	268	GCAGAAACCTGGAATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	327
Qy	361	GAATGGAATGTGTAATGTGACTTAAAGTCTGTGCGGATGATTTCTAACATGTGACA	420
Db	328	GAGCTTGACTGTGATGAAACCAATTTACGAGCTGTTCATGCGTTTCTTCAATGTGACT	387
Qy	421	TTACTGTCTTAAAGAAAACCAATTCACAGTCTTCCAGATTAAGTTTCAATCAATAC	480
Db	388	GCAATGTCACTTCAGTGAACCTTAATAAGAAAGTCTCTCTGATTTGCTTCAAGATTAAT	447
Qy	481	ACAAACCTTAAAGAAATATTTCTTCAGATTAATTCATTAACACATATCCAGAAAGCA	540
Db	448	CATGATCTTCAGAACTGTACCTGCAAAACATTAAGATTAATTCATTCATCTCATCTATCT	507
Qy	541	TTTTTGGATTAATGTAATCTGCAAAATATTATATCTCAACCAACTGATCAACACCTC	600
Db	508	TTCAAGAGACTGAATAGCCTTAATCACTTAATCTCACTCACTCACTCACTCACTCACT	567
Qy	601	AGACCTGGAATTTCAAAAGACTTACATCACTTAATCTGCTAATCTTGAATGACATCA	660
Db	568	AAGCGGGTGTGTTTGAAGATCTTCAACAGACTAGATGCTGATTAATGAAGATATAC	627
Qy	661	ATTAACCAATTTTCAACAGCTGTGTTTACGGGATTAATCTGTTTCTGCTATG	720
Db	628	CTCACTCGAATTTTCCCAACCAATTTATGAGACTTAATTTCTTAATCTTAACTCTG	687
Qy	721	GTTAATTAATCTTGAAGCTCTTCC---CAAGCAGATGTGTGCCAAATGCTCACTC	777
Db	688	ATGAATTAACGCTCAACCGCTTACCTGATTAACCTCTGTGCAACATGCGCAAGACTA	747
Qy	778	AAGTGGTGTGATTTGGAAGGCAATGAAATATCTCAAAATTTCAAGCTTCTGTG	837
Db	748	CATTGGCTGGAACCTTGAAGGCAACCATATCCATTAATTAAGAAATTTGACTTTATTTCC	807
Qy	838	TGCGATGCTCAGAGTGTGTTCTGCTAGAAATCAAAATGTTTGTCCAGAGAAG	897
Db	808	TGCAATTAATTTAACTGTTTATGATGATGAGAAAAAATAATTAATCACTTAATGAAAT	867
Qy	898	ACATTTTCTTCAATTAATAAATTTAGAGAATCTGTCTGACAAATGATTAACGAG	957
Db	868	ACTTTTGAACCTCTCAAGAACTGATGAATTTGATTAAGAAATTAATGAATGAAAT	927
Qy	958	CTATCACTCACTTTTAAAGACTTTCACAAACCTGAACCTGATCAATCAAT	1017
Db	928	CTTCAACCGCTTATTTCAAGAACTGGAAGGCTGTCAATTAATTTTCTTAATAT	987
Qy	1018	CCTTATGATCTTCAACAAAGACGTTTGAAGCTTAAACAACTTCACTGCTGAGAC	1077
Db	988	CCAATTCAGAAATTTCAAGAAACCAATTTGATTTATCTTGAACCTGATCACTGAC	1047
Qy	1078	CTGGAAGGATGAAGATTTCAATTAATTAACACAGATTTTCAACCACTGAAGAACTT	1137
Db	1048	CTAGAAAGGATTTGAATTTCAATATCCAAAGAGATTTTGAAGCTTCTTATGATCTC	1107
Qy	1138	TCTCAATTTATTTCAAAAGCTTGTGATCTGCTGATGCTCCCATGCTGCAATATGT	1197
Db	1108	TCTCAATTTATTTTAAGAAATTTCAAGTACTGTGGTATGACCACTATTTGCGAGCTGT	1167

Oy	1198	ATGCCCTTGAAGGACGGGATTTCTTCATTGGAGACCTCTGGGCTAAACAATTCCTCGA	125
Db	1168	AAACCAACACTGATGAAATTTCACTCTAGGAATCTCTGGGACGATTTACAGA	122
Oy	1258	ATATTGTCTGGGTAAAGCTTTCATTAACCTGCTTTGGAAATCTTTTGTCAATGGCAGT	131
Db	1228	GTAATTGCTGGGTGTATCTGCAGTTACCTGCTTTGGAAACATTTTGTCAATGGCAGT	128
Oy	1318	AGATTTTCATTAAAGCTGAAATACAACTCAACGTAATGTCCATCAAAATCCTTTGTTGT	137
Db	1288	CGACCTTATATCAGGCTGTAGAAACAGGCTGATAGCAATGTCATATCTCTGCTGCT	134
Oy	1378	GCTGATTCCTGATGGGTGTTTACTGTGTTCTTTGTGGCATTTTGCATATAAATACGGA	143
Db	1348	GCCGACTGCTTAATGGGAATATATTTATTCGTGATCGAAGGCTTTGACCTAAAGTTTGT	140
Oy	1438	GGGCAATATCAGAAATATGCTTGTCTGTGATGAGAGAGCTGACAGTCCGCTCATGGGG	149
Db	1408	GGAGAAATCAATTAACATGCGCAGCTGTGAGGAGAGAACTCATTTGTCAGCTTTAGGA	146
Oy	1498	TTCCGGGACATGCTGTCCACCGAAGCTCTGTTGCTACTGACCACTTGTGACTTTGGAG	155
Db	1468	TCTTTGGCATTTCTGTCCAGAGATATCAGTTTATCTGTTAACATTTCTGCAATTGGAA	152
Oy	1558	AAGTTCCTGTCATTTGCTTCCCTCACTTAACATTCGACCTGGAACACGACAGCTCA	161
Db	1528	AAATACATCTGCAATGTTCTATCTTTTGAATGTGAGACCTGGAACATGACAGAACAT	158
Oy	1618	GTCAATCCCTCATTTGCATCTGATGAGCGGATTTTAAATAGCTGTAAATCCATTTTGGAT	167
Db	1588	ACAGTTCGATATCTCATTTGGATATCTGGTTTAAATGAGGCTTCAATTCATTTGAGCAAT	164
Oy	1678	AAGATTAATTTTGGAACTTTTATGGAACAAATGAGATATGTTCCACTTATTAATGAC	173
Db	1648	AAGGAATTTTCAAAAACTACTATGGACCAATGAGATATGCTTCCCTCATTTCAAGA	170
Oy	1738	CAAAAGAAAGATTTTGGAAAGCAAAGGATTTCTCTTGGAAATTTTCTAGGTGTGAATTTG	179
Db	1708	GATACAGAAAGATTTGAGAGCCCAAGATTTTAAATGAGGCAATTTTCTTGGTATTAATTTG	176
Oy	1798	CTGGCTTTTCTCATATGTTGTGTTTCTATATTACTATATGTTCTGTTCCATTCAAAAAAC	185
Db	1768	GCCGACTTATATCATAGATTTTTCCTATGGAAGCATGTTTAAATGTATCATCAAAAT	182
Oy	1858	GCTTTCAGACCAAGAAATGAAGATGTTTGGAAAGAGGTGGCTGTGCAATCGT	191
Db	1828	GCCATTAACAGCAACTGAATACGGAATCAAGTTAAAAAAAGATGATTCCTGCCAAAGT	188
Oy	1918	TTCTTTTAAATAGTGTCTCTGATGCCATCTGCTGATTCCTGTATTTGTATTTAAATC	197
Db	1888	TTTTCTTTAATGATTTATCTATGATGATTAATCTGGATACCAATTTTGTATGTGAATTT	194
Oy	1978	CTTTCCTCTTCGCGGTGGAATATACAGACAATAGCACTCTGATAGTGAATTTTTC	203
Db	1948	CTTTACATGCTTCAAGTGAATACCAAGTATCATTAACCTCTTGGGTATGATTTTAAAT	200
Oy	2038	CTTCAATTAACAGTCTTTGAATCCATCTCTATACTCTCAACAACACTTTTAAAG	209
Db	2008	CTGCCATTTAAACAGTCTTTGAACCCCAATTCCTATACCTGACCAACAACACTTTAA	206
Oy	2098	GACAAATGGAACACGCTGTGACAACA---TCAGAGGAATCAATTTCAAAATTA	215
Db	2068	GAAATGATTCATCGGTTTTTGTGATTAATTAATTAACAAAGAAATCTATGACAGCAAGGT	212
Oy	2155	AAAAAAGTTATCTACATCAATGTTGTGTGATAGA	2189
Db	2128	CAGAAACATATGCTCATCATTCATCTGGGTGGA	2162

RESULT 14

AAZ25345

ID AAZ25345 standard; cDNA; 2467 BP.

AC	AAZ25345;
XX	
DT	20-DEC-1999 (first entry)
XX	
DE	Human LGR7 long form nucleotide sequence.
XX	
KW	Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy; extracellular leucine rich repeat region; mapping; identification; ss.
OS	Homo sapiens.
PN	WO948921-A1.
PD	
XX	30-SEP-1999.
PF	25-MAR-1999; 99WO-US06573.
PR	26-MAR-1998; 98US-0079501.
PA	(STRD) UNIV LELAND STANFORD JUNIOR. (ORGA) ORGANON NV.
XX	
FI	Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
DR	WPI; 1999-591074/50.
P-PDB:	AAI42170.
PT	New G-protein coupled receptors, useful for identifying their own ligands -
PS	Claim 4; Fig 3; 54pp; English.
CC	The present sequence encodes the human G-protein coupled receptor having extracellular leucine rich repeat regions, designated LGR7 long form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for the receptor. The polypeptides and/or polynucleotides are also useful for homologous or related genes, producing compositions that modulate the expression or function of the receptors, gene therapy, mapping functional regions of the receptors, studying associated physiological pathways, in vivo prophylactic and therapeutic purposes, as immunogens for producing antibodies, and for identifying biologically active agents. The polypeptides contain a G-protein coupled seven transmembrane region and a leucine rich repeat extracellular domain. These regions capture and facilitate optimal orientation of its ligand. The proteins are also expressed in diverse tissues.
SQ	Sequence 2467 BP; 747 A; 487 C; 474 G; 759 T; 0 other;
Query Match	32.5%; Score 735; DB 20; Length 2467;
Best Local Similarity	60.9%; Pred. No. 4,1e-186;
Matches 1663; Conservative	0; Mismatches 785; Indels 27; Gaps 3
Y	142 GGATATTTCCTCGGGGAAATCTTACCAAGTCTTAACCCGAGCTTTTCATGTGATGC 201
D	281 GGCTATTTCCTCGTGGGGAATCATCAAAGTCTTGCTCAGACTCTGACGTAAAGCT 340
Y	202 AAGATGACTGTGGGACGGGGCGGACGAAAGAAGTGTGTGACACTTAGTGATGGGCG 261
D	341 GTGGACGACTGGGGGAATCAGGCCCATGAGCAACTGTGGAGACAACAATGATGTCTC 400
Y	262 ACCAATTTTGGCAGCTGATGGAATGCTAAC-----GCGTG 300
D	401 ATGCAATTTGACAAATATTTTGCCAGTACTACAAAAGACTTCCCAATATCTTTGAG 460
Y	301 GCCTTAACACAGAGTGTCTTTTAAACAGTATCCAATGTCTGACTGCAAAGAACT 360
D	461 GCGAGAAACACCGAATGTTTGGTCGGTTCTGTGCCAGTGCATATGCTTTGCCAAGTCTG 520
Y	361 GAATGGAGATGTGTAATGTGTGACTTAAAGTGTGCGGATGATTTCTTAACAATGAGACA 420
D	521 GAGCTTGACTGTGATGAACAACATTTAGAGCTGTTCCATCGGTTCTTCAATATGTACT 580
Y	421 TTACGTCTCTTAAGAAAAACAAATTCACAGTCTTCCAGATTAAGTTTCAATAATAC 480

PI Hauenh AJW, Hsu SY, Liang S, Van Der Spek PJ;
 XX WPI; 1999-591074/50.
 DR P-PSDB; AAY42171.
 XX
 PT New G-protein coupled receptors, useful for identifying their own
 PT ligands -
 XX
 PS Claim 4; Fig 4; 54pp; English.
 XX
 CC The present sequence encodes the human G-protein coupled receptor
 CC having extracellular leucine rich repeat regions, designated IGR7 short
 CC form. The IGRs, IGR5 and IGR7 proteins are used to identify ligands for
 CC the receptor. The polypeptides and/or polynucleotides are also useful
 CC for homologous or related genes, producing compositions that modulate
 CC the expression or function of the receptors, gene therapy, mapping
 CC functional regions of the receptors, studying associated physiological
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
 CC for producing antibodies, and for identifying biologically active
 CC agents. The polypeptides contain a G-protein coupled seven
 CC transmembrane region and a leucine rich repeat extracellular domain.
 CC These regions capture and facilitate optimal orientation of its ligand.
 CC The proteins are also expressed in diverse tissues.

XX Sequence 3584 BP; 1124 A; 670 C; 647 G; 1139 T; 4 other;

Query Match 30.3%; Score 686.2; DB 20; Length 3584;
 Best Local Similarity 62.0%; Pred. No. 5.7e-173;
 Matches 1122; Conservative 0; Mismatches 683; Indels 6; Gaps 2;

QY 385 TTTAAGTCTGCGCATGATTTCTAACAATGACATTAATCTCTTAAAGAAAACAA 444
 DB 370 TTTAGAGTGTTCACGCGTTCCTCAATGATGACATGACATTCACAGGAACTTA 429
 QY 445 ATCCAGCTCTCCAGATAAGTTTTCATCAAAATACAAAACCTTAAAGATATTCTT 504
 DB 430 ATAAAGAGCTTCCTCGATGCTTCAAGATTATCATGATCTTCAAGAGCTGACCTG 489
 QY 505 CAGCATTAATGTCATTAGACATATTCAGAGAAAGCATTTTGGATTATGTAATCTGAA 564
 DB 490 CAAAACATTAAGATTATACATCCATCCATCTCATGCTTACAGAGACTGAATAGCTTACT 549
 QY 565 ATATTATATCTCAACACACATGTCATCAACCTCCAGACCTGGAAATTTCAAAACTTA 624
 DB 550 AACTGTAATCTAGTCATTAACAGAAATACCTTCGAAAGCGGGTGTGTAATCTT 609
 QY 625 CATCAGTAACTTGGCTAATTTCTAGATGACATCCAAATACAGAAATTTCAAGGCTTG 684
 DB 610 CACAGACTAGATGCTGATATTAAGATATACCTCAGTCAATTTCCCAACCAACA 669
 QY 665 TTTACGGATTAATTTCTGTTTTTCTGCTATGTTAATTAATTAATTAATTAATTA 744
 DB 670 TTTTATGACATTAATTTCTTATCTTATCTTATGCTGATGAATATATGCTCCACCTTTA 729
 QY 745 CC---CAGGAGATGTGCGCCCAATAGCTCACTCACTGAGTGGATTTGGAAGCAAT 801
 DB 730 CCTGATTAACCTCTGTCACACATGCGAAGACTCAATGCTGACCTTGAAGGCAAC 789
 QY 802 AGAATAAGTATCTCAAAATCTCAAGTTTCTGCTGCGATTTGCTCAGAGTCTGTT 861
 DB 790 CATATCCATTAATTAAGAAATTTGACTTTATTTCTGCAATTAATTAATTAATTA 849
 QY 862 CTGCTTGAATCAATTTGTTTGTTCAGAGAAAGCAATTTCTTCAATTAATAATTTA 921
 DB 850 ATGAGAAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 909
 QY 922 GGAGAACTGATCTGCTAGCAATAGATTAAGGAGCTATCACTCACCCTTTTAAAC 981
 DB 910 GATGATTTGATTTGAGAAATTAAGTAAATCTTCAACCCCTTAATTAATTAAGAC 969
 QY 982 TTGAAGCTTCTAACAAGCTGACCTGTCATCAATCTCTTATGATTTATCTTCAAGAAG 1041
 DB 970 CTGAAGAGCTGTCAATTAATTAATTTCTTCTATTAATCAATCAAAATTAACAGAAAC 1029

QY 1042 CAGTTGAAAGCTTTAAACAATTCAGTCTCTAGACCTGGAAAGATAGATTTCAAAAT 1101
 DB 1030 CAATTGATTAATCTTGCAAACTCAAGCTCTCAGCCTAGAAAGGATTAATTTCAAAAT 1089
 QY 1102 ATAAACACAGAAATGTTTCAACCAATGAAGAAATTTTCTCATTTATTTCAAAACCTT 1161
 DB 1090 ATCCAAAGAGAGTGTAGACCTTATGATATCTCTCATATATTTTAAAGAAATTC 1149
 QY 1162 CGATAGCTCTCTAGCTCCCATGTCGAAATATGTAATGACCTTGAAGGAGCATTTCT 1221
 DB 1150 CAGTACTGTGGTATGACCAACATGTTGCGAGCTTAAACCAACATGATGAAATTTCA 1209
 QY 1222 TCATTTAGAGACCTTTGGCTTAACAAATATCTCAGAAATTTGTCTGGTTATAGCTTTC 1281
 DB 1210 TCTCTAGAGAAATCTTGGCAAGCATTAATCAGAGAGATTTGTCTGGGTTGATCTGCA 1269
 QY 1282 ATTAACCTGCTTGAAGAAATTTTGTCAATGAGATGATTTTCAATTAAGCTGAAT 1341
 DB 1270 GTTACCTGCTTGGAAACATTTTGTCAATGAGATGAGACCTTATATCAGGCTGAGAAC 1329
 QY 1342 ACAACTACGCTATGTCATCAAAATCCTTGTGTGCTGATGCTGATGAGGCTTATAC 1401
 DB 1330 AAGCTGATGCTCAATGATTAATTTCTCTGCTGTGCGACCTTATATGGAATATAT 1389
 QY 1402 TTGTTCTTGTGTCATTTTGCATTAATAATACCGAGGCGATATCAGAAATATGCTTTC 1461
 DB 1390 TTATTTGATTCGAGAGCTTGAACCTTAAAGTTTGTGGAAGATTAATTAAGCATGCGCAG 1449
 QY 1462 CTGTGATGAGAGAGCTGACGCTGCTCATGAGGCTTCTGCGCATGCTTCCACCGAA 1521
 DB 1450 CTGTGATGAGAGATCTCATATGTCAGCTGTGAGATCTTGGCAATCTGTCCACAGAA 1509
 QY 1522 GTCTCTGTCTGCTACTGACCTTAATCTTGAAGAGTTCCTGGTCAATGCTTCTCCCC 1581
 DB 1510 GTATCAAGTTTATCTGTTAACTTTCTGACATGGAATAATCATCTGATTTCTATCT 1569
 QY 1582 TTCAGTAACATTCGACCTGGAAACGCGACGCTCACTCATCTTATGATCTGATG 1641
 DB 1570 TTTAGATGTGAGACCTGGAAATGCAAACTTACAGTTCTGATCTCATTTGGATT 1629
 QY 1642 GGGGATTTTAAATAGCTGTAATTCATTTTGAATAAGATTATTTTGAATCTTTAT 1701
 DB 1630 ACTGTTTTATAGGCTTTCATTTCAATGAGCAATTAAGAAATTTTCAAAACTACTAT 1689
 QY 1702 GGGAAATGAGATGATTTTCCACTTTATTAATGACCAACAGAAATATGGAAGCAA 1761
 DB 1690 GGCACCAATGAGATGCTCTCCCTTCTTCAATGACAAATGACAAAGTATGAGCCAG 1749
 QY 1762 GGGATTTCTCTGGAATTTTCTAGGTGGAATCTGTGCTTTCATCATTTGTTT 1821
 DB 1750 ATTATTCAGTGGCAATTTTCTTGATTAATTTGGCCGATTTATCATATGTTTTT 1809
 QY 1822 TCCATATTAATGATTTCTGCTCATTTCAAAAACCGCTTGCAGACCAAGAAATAGG 1881
 DB 1810 TCCATGGAACAGATTTTATAGTTCATCAAAAGTGCATTAACAGCAACGAAATACGG 1869
 QY 1882 AATGTTTTGGAAGAGAGTGGCTGTGCAAAATGTTTTCTTTTATATGATTTCTGAT 1941
 DB 1870 AATCAAGTAAAGAAAGATATCTTGCCAAAGCTTTTCTTATTAATTAATTAAT 1929
 QY 1942 GCAATCTGCTGATCTGTAATTTGATTAATAATCTTTTCCCTCTTCCGGGTGAATA 2001
 DB 1930 GCATTAATGCTGATACCAATTTTGTGATGAATTTCTTCACTGCTTCAAGTGAATA 1989
 QY 2002 CCAAGACAAATGATCTTCTGATAGTATTTTCTCCAGTAAACAGGCTTTGAT 2061
 DB 1990 CCAAGTACATTAATCTTGGGTAGATATTTTATCTGCCATTAACAGGCTTTGAT 2049
 QY 2062 CCAATCTCTATATCTTCAACCAACTTTTAAAGACAAGTGAACAGCTCTGAC 2121
 DB 2050 CCAATTTCTATATCTTGACCAAGAACATTTTAAAGAAATGATTAATCGGTTTGGAT 2109

Qy	2122	AAACA---TCAGAGAAATCAATTTCAAAATTA	AAAAAGTTTATCTACATCCATT	2178
Db	2110	AACTACAGACAAAGAAATCTATGACAGCAAGGT	CAGAAACATATGCTCCATCATTG	2169
Qy	2179	GTGTGGATAGA		2189
Db	2170	ATCTGGGTGGA		2180

Search completed: November 28, 2003, 15:51:45
 Job time : 595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 28, 2003, 15:41:35 ; Search time 722 Seconds
(without alignments)
10335.835 Million cell updates/sec

Title: US-09-928-175-1

Perfect score: 2265
Sequence: 1 atgattgttttcggtttc.....taatgaaccagtttcctag 2265

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2265	100.0	2265	10	US-09-928-175-1
2	2265	100.0	2265	14	US-10-223-668-1
3	2261.8	99.9	2261.8	14	US-10-223-735-2
4	2214	97.7	2214	11	US-09-965-536A-1
5	2111	93.2	2193	10	US-09-928-175-6
6	2060	90.9	2142	11	US-09-965-536A-5
7	1803	79.6	2049	10	US-09-928-175-11
8	1579.8	69.7	2539	14	US-10-223-735-4
9	1579	69.7	2214	10	US-09-928-175-19
10	1149	50.7	1149	10	US-09-928-175-4
11	1066.4	47.1	1068	12	US-10-321-807-17
12	995	43.9	1077	10	US-09-928-175-9
13	991	42.8	1101	10	US-09-928-175-16
14	825.8	36.5	1140	10	US-09-928-175-22
15	735	32.5	2274	14	US-10-225-567A-622
16	687	30.3	933	10	US-09-928-175-14

17	496.8	21.9	1804	10	US-09-895-686-10	Sequence 10, Appl
18	413	18.2	853	12	US-10-017-161-1325	Sequence 125, Ap
19	370	16.3	592	9	US-09-811-284-43	Sequence 43, Appl
20	321	14.2	321	11	US-09-930-112-1	Sequence 1, Appl
21	265.4	11.8	760	10	US-09-764-877-1029	Sequence 1029, Ap
22	232.4	10.3	252	12	US-10-029-386-16266	Sequence 16266, A
23	232.4	10.3	571	12	US-10-029-386-2566	Sequence 2566, Ap
24	232.4	10.3	576	9	US-09-811-284-100	Sequence 100, App
25	227.6	10.0	1126	14	US-10-313-542-255	Sequence 255, App
26	227	9.8	612	10	US-09-895-686-37	Sequence 37, Appl
27	221	9.8	1018	11	US-09-895-686-37	Sequence 65, Appl
28	185	8.2	562	10	US-09-895-686-38	Sequence 38, Appl
29	184.4	8.1	530	9	US-09-764-853-137	Sequence 137, App
30	184.4	8.1	530	10	US-09-764-853-986	Sequence 986, App
31	184.4	8.1	530	11	US-09-989-442-33	Sequence 33, Appl
32	184.4	8.1	530	12	US-09-764-886-18	Sequence 18, Appl
33	184.4	8.1	530	14	US-10-073-865-25	Sequence 25, Appl
34	184.4	8.1	530	14	US-10-103-213-89	Sequence 89, Appl
35	184.4	8.1	530	14	US-10-073-885-47	Sequence 47, Appl
36	183.2	8.1	1089	14	US-10-073-885-20	Sequence 20, Appl
37	173	7.6	466	10	US-09-895-686-34	Sequence 34, Appl
38	173	7.6	719	12	US-09-814-353-18321	Sequence 18321, A
39	167.4	7.4	1228	12	US-10-353-690-19	Sequence 19, Appl
40	150.4	6.6	677	12	US-09-814-353-5651	Sequence 5651, Ap
41	150.4	6.6	677	12	US-09-814-353-11937	Sequence 11937, A
42	113.8	5.0	1080	14	US-10-270-333-26	Sequence 26, Appl
43	112.8	5.0	1015	14	US-10-270-333-20	Sequence 20, Appl
44	81.2	3.6	1866	12	US-09-766-5118-22	Sequence 22, Appl
45	81.2	3.6	1866	12	US-10-188-495-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-09-928-175-1
Sequence 1, Application US/09928175
Patent No. US20020123618A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher J.
APPLICANT: Gong, Jianhua
APPLICANT: Daugherty, Betsy
APPLICANT: Rogers, No. US20020123618A1ma
TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
FILE REFERENCE: 00-1229
CURRENT APPLICATION NUMBER: US/09/928,175
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/224,455
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2265
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (2262)
NAME/KEY: sig_peptide
LOCATION: (1) .. (108)
US-09-928-175-1

Query Match 100.0%; Score 2265; DB 10; Length 2265;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGTTTTCGTTTAAACATCTCTTCAGCTCAAGTGAATGATGTTCTTT 60
DB 1 ATGATGTTTTCGTTTAAACATCTCTTCAGCTCAAGTGAATGATGTTCTTT 60
QY 61 CTACTGATTCATCGTTCGATCAATGTCAGAGTTTGCACGACCAAGTGCATG 120

Db 61 CTACTTCATTCATCGTTCGATCAANGTCAAGAATTTTGGACGACTCAAGGTAGATG 120
Qy 121 ATCACTCTTCATGCGCAAAAAGATATTTTCCCTGSGGAATCTTACCAAGTCTTACC 180
Db 121 ATCACTCTTCATGCGCAAAAAGATATTTTCCCTGSGGAATCTTACCAAGTCTTACC 180
Qy 181 CGAGCTTTTCACTGTGATGGCAAGAGTACGTGTGGGAGCGGGGCGAGCAAGAACTGT 240
Db 181 CGAGCTTTTCACTGTGATGGCAAGAGTACGTGTGGGAGCGGGGCGAGCAAGAACTGT 240
Qy 241 GGTGACACTAGTGGATGGCGACCATATTTTGGCACTGTGATGGAAATGTCTAACAGCTG 300
Db 241 GGTGACACTAGTGGATGGCGACCATATTTTGGCACTGTGATGGAAATGTCTAACAGCTG 300
Qy 301 GCGTTTAAACAAGAGTCTTTCTTAAACAGTATCTCACAATGCTGTGACTGCAAGAACT 360
Db 301 GCGTTTAAACAAGAGTCTTTCTTAAACAGTATCTCACAATGCTGTGACTGCAAGAACT 360
Qy 361 GAAATGGGAATGTGAAATGGGACTTAAAGTCTGTGCGAATGATTTCTTAACAATGTGACA 420
Db 361 GAAATGGGAATGTGAAATGGGACTTAAAGTCTGTGCGAATGATTTCTTAACAATGTGACA 420
Qy 421 TTAAGTCTCTTAAAGAAAAAACAATCCAGTCTTCCAGATTAAGTTTCTATCAATAC 480
Db 421 TTAAGTCTCTTAAAGAAAAAACAATCCAGTCTTCCAGATTAAGTTTCTATCAATAC 480
Qy 481 ACAAACCTTAAAGAAATTTTCTTCAAGCTTAAATGGCATTTAGACATATCCAGAAAGCA 540
Db 481 ACAAACCTTAAAGAAATTTTCTTCAAGCTTAAATGGCATTTAGACATATCCAGAAAGCA 540
Qy 541 TTTTGTGATTAATGTAATCTGCAATATTTATCTCAACCAACACTGATCAACCTC 600
Db 541 TTTTGTGATTAATGTAATCTGCAATATTTATCTCAACCAACACTGATCAACCTC 600
Qy 601 AGACCTGGAATATTCAAAGACTTACATCAGCTTAACTGGCTAAATCTAGATGACATCCA 660
Db 601 AGACCTGGAATATTCAAAGACTTACATCAGCTTAACTGGCTAAATCTAGATGACATCCA 660
Qy 661 ATAAACGAATTTTCAAGCGCTTGTGTTTACGGGATTTAAATTCCTGTGTTTCCGTCTAGT 720
Db 661 ATAAACGAATTTTCAAGCGCTTGTGTTTACGGGATTTAAATTCCTGTGTTTCCGTCTAGT 720
Qy 721 GTTATATACTACTAGAAAGCTCTTCCCAAGCAGATGTGTGCCAAATGCTCAACTCAAC 780
Db 721 GTTATATACTACTAGAAAGCTCTTCCCAAGCAGATGTGTGCCAAATGCTCAACTCAAC 780
Qy 781 TGGGTGATTTTGAAGGCAATAGATAAATGATCTCAAAATCTTACGTTTCTGTCTGTC 840
Db 781 TGGGTGATTTTGAAGGCAATAGATAAATGATCTCAAAATCTTACGTTTCTGTCTGTC 840
Qy 841 GATTCGCTCAAGTGTCTTCTGCTGATGAAATCAAAATGGTTTGTCCAGAGAGCA 900
Db 841 GATTCGCTCAAGTGTCTTCTGCTGATGAAATCAAAATGGTTTGTCCAGAGAGCA 900
Qy 901 TTTTCTTCATTTAAAAAATTTAGAGAACTGATCTGTCTAGCAATAGATAACGAGCTA 960
Db 901 TTTTCTTCATTTAAAAAATTTAGAGAACTGATCTGTCTAGCAATAGATAACGAGCTA 960
Qy 961 TCACCTCACTTTTAAAGACTTGAAGCTTCTTACAAAAGCTGGAACCTGTCTCAATCTCT 1020
Db 961 TCACCTCACTTTTAAAGACTTGAAGCTTCTTACAAAAGCTGGAACCTGTCTCAATCTCT 1020
Qy 1021 CTTATGTATCTTCAAGAGCAAGTTTGAAGTCTTAAACAACCTTCACTCTAGACCTG 1080
Db 1021 CTTATGTATCTTCAAGAGCAAGTTTGAAGTCTTAAACAACCTTCACTCTAGACCTG 1080
Qy 1081 GAAAAGATAGAGATTCCAATATATAACAACAAGATGTTTCAACCATGAAATCTTCTCT 1140
Db 1081 GAAAAGATAGAGATTCCAATATATAACAACAAGATGTTTCAACCATGAAATCTTCTCT 1140
Qy 1141 CACATTTATTTCAAAAACCTTGCATCTGCTCTATGCTCCCATGTCGAATATGATG 1200
Db 1141 CACATTTATTTCAAAAACCTTGCATCTGCTCTATGCTCCCATGTCGAATATGATG 1200

Qy 1201 CCCTTGAAGGAGCGCAATTTCTTCAATTTAGAGACCTTGGCTTAAACAATATCCGACAGATA 1260
Db 1201 CCCTTGAAGGAGCGCAATTTCTTCAATTTAGAGACCTTGGCTTAAACAATATCCGACAGATA 1260
Qy 1261 TTTGTCTGGGTTTATAGCTTTCATTAACCTGCTTGTGAAATCTTTTGTCTATTGGCATGAG 1320
Db 1261 TTTGTCTGGGTTTATAGCTTTCATTAACCTGCTTGTGAAATCTTTTGTCTATTGGCATGAG 1320
Qy 1321 TCTTTCATTTAAAGCTGAAAAATACACTCAAGCTATGCTCATCAAAATCTTTGTGTGCT 1380
Db 1321 TCTTTCATTTAAAGCTGAAAAATACACTCAAGCTATGCTCATCAAAATCTTTGTGTGCT 1380
Qy 1381 GATTCGCTGATGGGCTTTTACTTGTCTTGTGGAATTTTGGATTAATAATACGAGGG 1440
Db 1381 GATTCGCTGATGGGCTTTTACTTGTCTTGTGGAATTTTGGATTAATAATACGAGGG 1440
Qy 1441 CAGTATCAGAAAGTATGCTTGTGCTGATGAGAGAGCTGCAAGTCCGCTCATGGGGTTC 1500
Db 1441 CAGTATCAGAAAGTATGCTTGTGCTGATGAGAGAGCTGCAAGTCCGCTCATGGGGTTC 1500
Qy 1501 CTGGCAATGCTGTCCACCGAAGTCTGTCTGTCTACTGACCTTACTTGTGAGAG 1560
Db 1501 CTGGCAATGCTGTCCACCGAAGTCTGTCTGTCTACTGACCTTACTTGTGAGAG 1560
Qy 1561 TTTCTGTGATGCTTCCCTTCACTTAACTTCACTGAAATTCCTTGTGAGAG 1620
Db 1561 TTTCTGTGATGCTTCCCTTCACTTAACTTCACTGAAATTCCTTGTGAGAG 1620
Qy 1621 ATCTCATTTTGCATCTGATGAGGAGGATTTTAAATAGCTGTAATTCATTTTGAATTAAG 1680
Db 1621 ATCTCATTTTGCATCTGATGAGGAGGATTTTAAATAGCTGTAATTCATTTTGAATTAAG 1680
Qy 1681 GATTTATTTTGGAACTTTTATGAGGAAAAATGAGATGTTTCCACTTATATATGACCA 1740
Db 1681 GATTTATTTTGGAACTTTTATGAGGAAAAATGAGATGTTTCCACTTATATATGACCA 1740
Qy 1741 ACAGAAATATTTGGAAGCAAAAGGATTTCTTGGAAATTTTCTTGGTGTGAACTTGTCTG 1800
Db 1741 ACAGAAATATTTGGAAGCAAAAGGATTTCTTGGAAATTTTCTTGGTGTGAACTTGTCTG 1800
Qy 1801 GCTTTTCTCATATGTTGTTTCTTATATTAATGATGTTCTGTTCATTCATAAAGCCGC 1860
Db 1801 GCTTTTCTCATATGTTGTTTCTTATATTAATGATGTTCTGTTCATTCATAAAGCCGC 1860
Qy 1861 TTGCAAGCCACAAGATGAGAAATGTTTGTGGAAGAGGTGGCTGTGGAATCGTTTC 1920
Db 1861 TTGCAAGCCACAAGATGAGAAATGTTTGTGGAAGAGGTGGCTGTGGAATCGTTTC 1920
Qy 1921 TTTTATATAGTGTCTCTGATGCAATCTGCTGATTCCTGATTTGTGATTAATTCCTT 1980
Db 1921 TTTTATATAGTGTCTCTGATGCAATCTGCTGATTCCTGATTTGTGATTAATTCCTT 1980
Qy 1981 TCCCTCTTCCGGGTGGAATATCAAGACCAATGACTTCTGTGATGATTTTTCCTT 2040
Db 1981 TCCCTCTTCCGGGTGGAATATCAAGACCAATGACTTCTGTGATGATTTTTCCTT 2040
Qy 2041 CCAAGTTAAGAGTGTGATGCAATCTGATATGCTGCAACAACCACTTTTAAAGAC 2100
Db 2041 CCAAGTTAAGAGTGTGATGCAATCTGATATGCTGCAACAACCACTTTTAAAGAC 2100
Qy 2101 AAGTTGAAACAGCTGTGCAACAACATCAGAGGAATCAATTTTCAAAATTTAAAAA 2160
Db 2101 AAGTTGAAACAGCTGTGCAACAACATCAGAGGAATCAATTTTCAAAATTTAAAAA 2160
Qy 2161 AGTTTATCTATCTATGTTGTGATGAGAGACTCCCTTCCCTGAAACCTTGGGCTTTG 2220
Db 2161 AGTTTATCTATCTATGTTGTGATGAGAGACTCCCTTCCCTGAAACCTTGGGCTTTG 2220
Qy 2221 AACAAATATACACTTGGAGACAGTATATGAAACCAAGTTTCTTGA 2265
Db 2221 AACAAATATACACTTGGAGACAGTATATGAAACCAAGTTTCTTGA 2265

RESULT 2
US-10-222-668-1
Sequence 1, Application US/10222668
Publication No. US2003088884A1
GENERAL INFORMATION:
APPLICANT: Heueh, Aaron
TITLE OF INVENTION: Mammalian Relaxin Receptor
FILE REFERENCE: STAN-239 MO
CURRENT APPLICATION NUMBER: US/10/222,668
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: 60/313,259
PRIOR FILING DATE: 2002-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1
LENGTH: 2838
TYPE: DNA
ORGANISM: H. sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (107)...(2369)
US-10-222-668-1

Query Match 100.0%; Score 2265; DB 14; Length 2838;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGTTTTCTGCTTTTAAACATCTCTTCAAGCTTCAATGATTAACAATGTTCTT 60
DB 107 ATGATGTTTTCTGCTTTTAAACATCTCTTCAAGCTTCAATGATTAACAATGTTCTT 166
QY 61 CTACTCATTTTCATGCTTCTGATCTCATGTCAAAGATTTTGCACTGACCTCAAGTACATG 120
DB 167 CTACTCATTTTCATGCTTCTGATCTCATGTCAAAGATTTTGCACTGACCTCAAGTACATG 226
QY 121 ATCACTCTCTTCATGCAAAAAGATATTTTCCCTGTGGGAATCTTACCAAGTCTTACC 180
DB 227 ATCACTCTCTTCATGCAAAAAGATATTTTCCCTGTGGGAATCTTACCAAGTCTTACC 286
QY 181 CGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 240
DB 287 CGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAACGGGGCGGAGCAAGAGAACTGT 346
QY 241 GGTGACACTAGTGTGGGCGACCATATTTGSCACAGTGCATGGAATGCTAACAGCGTG 300
DB 347 GGTGACACTAGTGTGGGCGACCATATTTGSCACAGTGCATGGAATGCTAACAGCGTG 406
QY 301 GCTTTAACACAGAGTGTCTTTCTAAACAGTATTCACAAATGCTGTGACGTCAAGAACT 360
DB 407 GCTTTAACACAGAGTGTCTTTCTAAACAGTATTCACAAATGCTGTGACGTCAAGAACT 466
QY 361 GAATGGAATGTGTAAATGTGTAAAGTCTGTGCGCATATTTCTTAAACATGTGACA 420
DB 467 GAATGGAATGTGTAAATGTGTAAAGTCTGTGCGCATATTTCTTAAACATGTGACA 526
QY 421 TTACTGTCTTAAAGAAAACAAATCCACAGTCTTCCAGATTAAGTTTCTATCAATAC 480
DB 527 TTACTGTCTTAAAGAAAACAAATCCACAGTCTTCCAGATTAAGTTTCTATCAATAC 586
QY 481 ACAAAACCTTAAAGATATTTCTTACAGTAAATGCTATTGACATATTCACAGAAAGCA 540
DB 587 ACAAAACCTTAAAGATATTTCTTACAGTAAATGCTATTGACATATTCACAGAAAGCA 646
QY 541 TTTTGTGATTTATGTATCTGCAATATATATTTTCAACACACATGCTCAACCCCTC 600
DB 647 TTTTGTGATTTATGTATCTGCAATATATATTTTCAACACACATGCTCAACCCCTC 706
QY 601 AGACCTGGAATATTTCAAGACTTACATCAGTAACTTGGCTAATTTAGATGACATCA 660
DB 707 AGACCTGGAATATTTCAAGACTTACATCAGTAACTTGGCTAATTTAGATGACATCA 766

QY 661 ATACAGAAATTTTCAACAGCGCTTTTACGGGATTAATTCCTGTTTTCTGTCTATG 720
DB 767 ATACAGAAATTTTCAACAGCGCTTTTACGGGATTAATTCCTGTTTTCTGTCTATG 826
QY 721 GTTATATATCTTATGAGCTCTTCCCAAGAGTGTGCGCAATGCTCAACTCAAC 780
DB 827 GTTATATATCTTATGAGCTCTTCCCAAGAGTGTGCGCAATGCTCAACTCAAC 886
QY 781 TGGGTGATTTGGAAGCAATATGAAATATCTCAAAATTTAGGTTTCTGTGCGC 840
DB 887 TGGGTGATTTGGAAGCAATATGAAATATCTCAAAATTTAGGTTTCTGTGCGC 946
QY 841 GATTGCTCACAGTGTCTTCTGCTTCAAGATCAATATGTTTGTGTCAGAGAGCA 900
DB 947 GATTGCTCACAGTGTCTTCTGCTTCAAGATCAATATGTTTGTGTCAGAGAGCA 1006
QY 901 TTTTCTTCAATTAATAATTTTATGAGAACTGATCTGTCTAGCATATGATTAACGAGCTA 960
DB 1007 TTTTCTTCAATTAATAATTTTATGAGAACTGATCTGTCTAGCATATGATTAACGAGCTA 1066
QY 961 TCACCTCACCTTTTAAAGACTTGAAGCTTCAAAAAGCTGAACCTGTCAATCCATCCT 1020
DB 1067 TCACCTCACCTTTTAAAGACTTGAAGCTTCAAAAAGCTGAACCTGTCAATCCATCCT 1126
QY 1021 CTATATATCTTCAACAAACCAAGTTGAAGCTTAAACAATCTCACTCTAGACCTG 1080
DB 1127 CTATATATCTTCAACAAACCAAGTTGAAGCTTAAACAATCTCACTCTAGACCTG 1186
QY 1081 GAAGGATAGAGATTTCCAAATATTAACACAGAAATGTTTCAACCCATGAGAAATCTTCT 1140
DB 1187 GAAGGATAGAGATTTCCAAATATTAACACAGAAATGTTTCAACCCATGAGAAATCTTCT 1246
QY 1141 CAGATTTATTTCAAAAACCTTCAAGTACGCTCTATGCTCCCAATGCGAATATGATG 1200
DB 1247 CAGATTTATTTCAAAAACCTTCAAGTACGCTCTATGCTCCCAATGCGAATATGATG 1306
QY 1201 CCTTTGACGAGCGGCAATTTCTTCAATTTGAGGACCTTTGAGTAAATATCTCAGAAAT 1260
DB 1307 CCTTTGACGAGCGGCAATTTCTTCAATTTGAGGACCTTTGAGTAAATATCTCAGAAAT 1366
QY 1261 TTTGTCTGGGTATATAGCTTTCATTAACCTGCTTTGGAATCTTTTGTGCAATGGATGGA 1320
DB 1367 TTTGTCTGGGTATATAGCTTTCATTAACCTGCTTTGGAATCTTTTGTGCAATGGATGGA 1426
QY 1321 TCTTTCAATTAAGCTGAATAATACAACTCAACCTATGTCATCAATAATCTTTGTGTGCT 1380
DB 1427 TCTTTCAATTAAGCTGAATAATACAACTCAACCTATGTCATCAATAATCTTTGTGTGCT 1486
QY 1381 GATTGCTGATGGGTGTTTACTGTCTTTGTGGCATTTTGCATTAATAAATACGAGGG 1440
DB 1487 GATTGCTGATGGGTGTTTACTGTCTTTGTGGCATTTTGCATTAATAAATACGAGGG 1546
QY 1441 CAGATTCAGAGTATGCTTCTGTGATGAGAGCGGACAGTGGCGGCTCATAGGGGTTC 1500
DB 1547 CAGATTCAGAGTATGCTTCTGTGATGAGAGCGGACAGTGGCGGCTCATAGGGGTTC 1606
QY 1501 CTGGCATGCTGTCCACCGAAGTCTGTGTCTGCTAATGACCTTACTGATCTTTGGAGAAG 1560
DB 1607 CTGGCATGCTGTCCACCGAAGTCTGTGTCTGCTAATGACCTTACTGATCTTTGGAGAAG 1666
QY 1561 TTCTGTGATATGTTCTTCCCTTCAATACATTTGACCTGAAAACGGGACGCTCACTC 1620
DB 1667 TTCTGTGATATGTTCTTCCCTTCAATACATTTGACCTGAAAACGGGACGCTCACTC 1726
QY 1621 ATTCCTATTTGATCTGATGAGCGGATTTTATATAGCTATATTCATTTTGGAAATAG 1680
DB 1727 ATTCCTATTTGATCTGATGAGCGGATTTTATATAGCTATATTCATTTTGGAAATAG 1786
QY 1681 GATTATTTTGGAAAATTTTATGAGAAAATGAGATGTTTCCCATTAATATATGACCA 1740
DB 1787 GATTATTTTGGAAAATTTTATGAGAAAATGAGATGTTTCCCATTAATATATGACCA 1846
QY 1741 ACAGAGATATTTGAGAGCAAGGATATTTCTTGTGAATTTTCTATGATGTGAACTTGGCTG 1800

Db 1847 ACGAAGATATTTGGAAGCAAGGATATTCCTTGGAAATTTCTAGGTGGAACCTTGCG 1906
Qy 1801 GCTTTTCTCATCTTGTGTTTCTTATATTAATGTTCTGTTCCATTCAAAAACCGCC 1860
Db 1907 GCTTTTCTCATCTTGTGTTTCTTATATTAATGTTCTGTTCCATTCAAAAACCGCC 1966
Qy 1861 TTGCAGACCAAGAGTAAGAAATGTTTGAAGAGGAGGCTGTGCAAAATCGTTTC 1920
Db 1967 TTGCAGACCAAGAGTAAGAAATGTTTGAAGAGGAGGCTGTGCAAAATCGTTTC 2026
Qy 1921 TTTTATATAGTGTCTCTGATGCAATCTGCTGATTCCTGTAATTTGATTAATTCCTT 1980
Db 2027 TTTTATATAGTGTCTCTGATGCAATCTGCTGATTCCTGTAATTTGATTAATTCCTT 2086
Qy 1981 TCCCTCTCCGGGTGGAATATCAAGACAAATGATCTTCTGATGATGATTTTTCCTT 2040
Db 2087 TCCCTCTCCGGGTGGAATATCAAGACAAATGATGATGATGATTTTTCCTT 2146
Qy 2041 CCAGTTAAGAGTCTTGAATCCAAATCTATACCTCAACCACTTTTAAAGAC 2100
Db 2147 CCAGTTAAGAGTCTTGAATCCAAATCTATACCTCAACCACTTTTAAAGAC 2206
Qy 2101 AAGTTGAAACAGCTGTGCAAAACATCAGAGGAATCAATTTTCAAAATTAATAAAAAA 2160
Db 2207 AAGTTGAAACAGCTGTGCAAAACATCAGAGGAATCAATTTTCAAAATTAATAAAAAA 2266
Qy 2161 AGTTTATCTACATCCATGTTGTGATGATGAGACCTCTCTCCGTAACCTTGGGCTTTC 2220
Db 2267 AGTTTATCTACATCCATGTTGTGATGATGAGACCTCTCTCCGTAACCTTGGGCTTTC 2326
Qy 2221 AACAAATAACACTTGAGAGACAGTATATGAAACCAATTTCTAG 2265
Db 2327 AACAAATAACACTTGAGAGACAGTATATGAAACCAATTTCTAG 2371

RESULT 3
US-10-229-735-2
; Sequence 2, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
; APPLICANT: Baylor College of Medicine
; APPLICANT: Agoulnik, Alexander I.
; TITLE OF INVENTION: The GREAT Gene and Protein
; FILE REFERENCE: 7572/73263
; CURRENT APPLICATION NUMBER: US/10/229, 735
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/315, 696
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/351, 432
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2436
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-229-735-2

Query Match 99.9%; Score 2261.8; DB 14; Length 2436;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2263; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGATGTTTTTCTGTTTTTAAACATCTCTGAGCCTCAGATTGATTACATGTTCTT 60
Db 71 ATGATGTTTTTCTGTTTTTAAACATCTCTGAGCCTCAGATTGATTACATGTTCTT 130
Qy 61 CTACTTCATTTTCATGTTCTGATCAATGTCAAGATTTTGCAGTCAAGAGTAGCATG 120
Db 131 CTACTTCATTTTCATGTTCTGATCAATGTCAAGATTTTGCAGTCAAGAGTAGCATG 190
Qy 121 ATGCTCTTATGCGCAAAAGATATTTTCCCTGTGGATCTTACCAAGTGTTCACC 180

Db 191 ATCACTCCTTACGCCAAAAAGATATTTTCCCTGTGGAAATCTTACCAAGTCTTACCC 250
Qy 181 CGAGCTTTTCACTGTGATGAGCAAGATGATCTGTGGGAACGGGGCGGACGAAGGAATCT 240
Db 251 CGAGCTTTTCACTGTGATGAGCAAGATGATCTGTGGGAACGGGGCGGACGAAGGAATCT 310
Qy 241 GGTGACATGATGATGAGGCGACCAATTTTGGCAAGTGTGATGAAATGCTAACAGCTG 300
Db 311 GGTGACATGATGATGAGGCGACCAATTTTGGCAAGTGTGATGAAATGCTAACAGCTG 370
Qy 301 GCTTTAACAAGAGTCTTTCTTAAACAGTATCCCAATGCTGTGACCTGCAAAAGAACT 360
Db 371 GCTTTAACAAGAGTCTTTCTTAAACAGTATCCCAATGCTGTGACCTGCAAAAGAACT 430
Qy 361 GAATGGAATGATGATGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 420
Db 431 GAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
Qy 421 TTAAGTCTCTTAAAGAAAAAATCCACAGTCTTCCAGATTAAGTTTCAATCAATAC 480
Db 491 TTAAGTCTCTTAAAGAAAAAATCCACAGTCTTCCAGATTAAGTTTCAATCAATAC 550
Qy 481 ACAAACCTTAAAGATATTTCTTCAAGCAATATGATTAACAATCCAGAAAGCA 540
Db 551 ACAAACCTTAAAGATATTTCTTCAAGCAATATGATTAACAATCCAGAAAGCA 610
Qy 541 TTTTGTGATTAATGATATGCAAAATATATATCTCAACCAACCTGATCAACCTCTC 600
Db 611 TTTTGTGATTAATGATATGCAAAATATATATCTCAACCAACCTGATCAACCTCTC 670
Qy 601 AAGCTGGAATATTCAGAGCTTACATCAGCTTACCTGCTAATTTAGATGCAATCCA 660
Db 671 AAGCTGGAATATTCAGAGCTTACATCAGCTTACCTGCTAATTTAGATGCAATCCA 730
Qy 661 ATTAACAGAAATTCAGAGCTTACATCAGCTTACCTGCTAATTTAGATGCAATCCA 720
Db 731 ATTAACAGAAATTCAGAGCTTACATCAGCTTACCTGCTAATTTAGATGCAATCCA 790
Qy 721 GTTAATTAATCTTAAAGCTCTTCCCAAGCAATGTCGCAAAATGCTCAACTCAAC 780
Db 791 GTTAATTAATCTTAAAGCTCTTCCCAAGCAATGTCGCAAAATGCTCAACTCAAC 850
Qy 781 TGGGTGATTTGGAAGCAATGATTAAGTATCTCAAAATTTCTGCTGTGTGTC 840
Db 851 TGGGTGATTTGGAAGCAATGATTAAGTATCTCAAAATTTCTGCTGTGTGTC 910
Qy 841 GATTCCTCAAGTGTGTTTCTGCTTAAGAAATCAATTTGTTTGTCCAGAGGAGCA 900
Db 911 GATTCCTCAAGTGTGTTTCTGCTTAAGAAATCAATTTGTTTGTCCAGAGGAGCA 970
Qy 901 TTTTCTTCAATTAATAATTTAGAGCAATGATCTGTCTAGCAATAGATTAACGAGCTA 960
Db 971 TTTTCTTCAATTAATAATTTAGAGCAATGATCTGTCTAGCAATAGATTAACGAGCTA 1030
Qy 961 TCACTCACTTTTAAAGCTTGAAGCTTCTAAGAAAGCTGAACCTGTATCAATCTCT 1020
Db 1031 TCACTCACTTTTAAAGCTTGAAGCTTCTAAGAAAGCTGAACCTGTATCAATCTCT 1090
Qy 1021 CTATGATCTTCAAGAAACAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTT 1080
Db 1091 CTATGATCTTCAAGAAACAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTT 1150
Qy 1081 GAAAGATGAGATTCGAATATTAACACAGATGTTTCAACCAATGAAGAAATCTTCT 1140
Db 1151 GAAAGATGAGATTCGAATATTAACACAGATGTTTCAACCAATGAAGAAATCTTCT 1210
Qy 1141 GAAATTTATTTCAAAACCTTGTGATCTGTCTTAAGTCTCCCAATGCGAATATGATG 1200
Db 1211 GAAATTTATTTCAAAACCTTGTGATCTGTCTTAAGTCTCCCAATGCGAATATGATG 1270
Qy 1201 CCGTTGACGAGCAATTTCTTCAATTTAGAGCACTCTTGGCTTAACAATATCTCAAGATA 1260
Db 1271 CCGTTGACGAGCAATTTCTTCAATTTAGAGCACTCTTGGCTTAACAATATCTCAAGATA 1330


```

Db      601 GACATCCAAATACAGAAATTCACAGGCGTTGTTAGGGATTAATTCCTGTTTTTC 660
Qy      712 CTGTCTATGTTAAATPACTACTTAAAGCTCTTCCCAAGCAGATGTGSCCAATTCGCT 771
Db      661 CTGTCTATGTTAAATPACTACTTAAAGCTCTTCCCAAGCAGATGTGSCCAATTCGCT 720
Qy      772 CAACCTCACTGGGTGGATTTGGAAAGGCATAGAAATTAAGTATCTCACAAAATCTAGCTT 831
Db      721 CAACCTCACTGGGTGGATTTGGAAAGGCATAGAAATTAAGTATCTCACAAAATCTAGCTT 780
Qy      832 CTGTCTATGTTAAATPACTACTTAAAGCTCTTCCCAAGCAGATGTGSCCAATTCGCT 891
Db      781 CTGTCTATGTTAAATPACTACTTAAAGCTCTTCCCAAGCAGATGTGSCCAATTCGCT 840
Qy      892 GAGAAACATTTTCTTCAATTAATAAATTTAGAGAACTGATCTGTCTGCAATACGATA 951
Db      841 GAGAAACATTTTCTTCAATTAATAAATTTAGAGAACTGATCTGTCTGCAATACGATA 900
Qy      952 AGGAGCTATCACTCCCTTTTAAAGCTTGAACCTCTCAAAAGCTGAACCTGTCA 1011
Db      901 AGGAGCTATCACTCCCTTTTAAAGCTTGAACCTCTCAAAAGCTGAACCTGTCA 960
Qy      1012 TCCATCTCTTATGTATCTTCAAGAAACAGTTTGAAGCTTAAAGCACTTCACTCT 1071
Db      961 TCCATCTCTTATGTATCTTCAAGAAACAGTTTGAAGCTTAAAGCACTTCACTCT 1020
Qy      1072 CTAGACCTGGAAAGATAGAGATTCGAATATTAACACAGATGTTTCAACCCATGAAG 1131
Db      1021 CTAGACCTGGAAAGATAGAGATTCGAATATTAACACAGATGTTTCAACCCATGAAG 1080
Qy      1132 AATCTTCTCAATTTATTTCAAAAATTGATGATGCTCTATGCTCCCATGTCGA 1191
Db      1081 AATCTTCTCAATTTATTTCAAAAATTGATGATGCTCTATGCTCCCATGTCGA 1140
Qy      1192 ATATGATGCTCTTGAAGCAGGCACTTCTTCAATTTGAGAGCTCTTGAACCAATATC 1251
Db      1141 ATATGATGCTCTTGAAGCAGGCACTTCTTCAATTTGAGAGCTCTTGAACCAATATC 1200
Qy      1252 CTGAGAAATATTTCTGAGTTAATAGCTTCAATTAACCTGCTTGAAGCTTCTTGTCAAT 1311
Db      1201 CTGAGAAATATTTCTGAGTTAATAGCTTCAATTAACCTGCTTGAAGCTTCTTGTCAAT 1260
Qy      1312 GGCATGAGATCTTCAATTAAGCTGAATAATACAACTGATGCTATGCTCAAAATCTT 1371
Db      1261 GGCATGAGATCTTCAATTAAGCTGAATAATACAACTGATGCTATGCTCAAAATCTT 1320
Qy      1372 TGTGTCGATGATGCTGATGAGGTGTTACTGTTCTTGTGTCATTTGATATATAA 1431
Db      1321 TGTGTCGATGATGCTGATGAGGTGTTACTGTTCTTGTGTCATTTGATATATAA 1380
Qy      1432 TACCGAGGAGATATGAGATATGCTTCTGTCGATGAGAGCGTCGATGCTGCTC 1491
Db      1381 TACCGAGGAGATATGAGATATGCTTCTGTCGATGAGAGCGTCGATGCTGCTC 1440
Qy      1492 ATGGGGTTCCTGCGCATGCTGTCCACGGAAGTCTGTTGTCATGACCTAATTGACT 1551
Db      1441 ATGGGGTTCCTGCGCATGCTGTCCACGGAAGTCTGTTGTCATGACCTAATTGACT 1500
Qy      1552 TTGGAGAAAGTTCTGTGTCATTTCTTCCCTTCAATTAACCTGATGCTGGAAGCGGAG 1611
Db      1501 TTGGAGAAAGTTCTGTGTCATTTCTTCCCTTCAATTAACCTGATGCTGGAAGCGGAG 1560
Qy      1612 ACCCTGATCATCTCCATTTGATGATGAGCGGGAATTTTAAATAGCTGATATTCATTT 1671
Db      1561 ACCCTGATCATCTCCATTTGATGATGAGCGGGAATTTTAAATAGCTGATATTCATTT 1620
Qy      1672 TGAATTAAGATTAATTTTGAACCTTTTATGGAAGAAATGAGATGTTTCCACTTAT 1731
Db      1621 TGAATTAAGATTAATTTTGAACCTTTTATGGAAGAAATGAGATGTTTCCACTTAT 1680
Qy      1732 TATATACCAACAGAAAGATTTGGAAGCAAGGATTTCTCTGGAATTTTCTTAGGTGTG 1791
Db      1681 TATATACCAACAGAAAGATTTGGAAGCAAGGATTTCTCTGGAATTTTCTTAGGTGTG 1740

```

```

Qy      1792 AACTGTCGCTTTTCTCATGATTTGTTTTCCTATATTAATGTTCTGTTCCATTCAA 1851
Db      1741 AACTGTCGCTTTTCTCATGATTTGTTTTCCTATATTAATGTTCTGTTCCATTCAA 1800
Qy      1852 AAAACCGCTTTCAGACACACAGAAATGAAATGTTTGGAAAGAGAGTGGCTGTCGA 1911
Db      1801 AAAACCGCTTTCAGACACACAGAAATGAAATGTTTGGAAAGAGAGTGGCTGTCGA 1860
Qy      1912 AATGCTTCTTTTATATGTTTCTCTGATGTCATCTGCTGATTTCTGATTTGACTT 1971
Db      1861 AATGCTTCTTTTATATGTTTCTCTGATGTCATCTGCTGATTTCTGATTTGACTT 1920
Qy      1972 AAAATCTTCCCTTCCTGCGGAGTGAATATACAGACAAATGATCTTCTGATAGTAT 2031
Db      1921 AAAATCTTCCCTTCCTGCGGAGTGAATATACAGACAAATGATCTTCTGATAGTAT 1980
Qy      2032 TTTTCTTCTCCAGTTTACAGTGTGTTGAATCCATCTCTATCTCTGACAAACCACTT 2091
Db      1981 TTTTCTTCTCCAGTTTACAGTGTGTTGAATCCATCTCTATCTCTGACAAACCACTT 2040
Qy      2092 TTTTAAAGCAAGTTGAAAACAGCTGCTGCAAAACATTCAGAGAAATCAATTTTCAAAAT 2151
Db      2041 TTTTAAAGCAAGTTGAAAACAGCTGCTGCAAAACATTCAGAGAAATCAATTTTCAAAAT 2100
Qy      2152 AAAAAAAGTTTATCTACATCCATTTGATGATAGAGAGTCTCTCTCCCTGAAACTT 2211
Db      2101 AAAAAAAGTTTATCTACATCCATTTGATGATAGAGAGTCTCTCTCCCTGAAACTT 2160
Qy      2212 GGGGTTTGAACAAATATTAACCTTGAAGCAGTATATGAAAACAGTTTCTAG 2265
Db      2161 GGGGTTTGAACAAATATTAACCTTGAAGCAGTATATGAAAACAGTTTCTAG 2214

```

```

RESULT 5
US-09-928-175-6
; Sequence 6, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszcy, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2193
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2190)
; NAME/KEY: s19_peptide
; LOCATION: (1)..(108)
US-09-928-175-6

```

```

Query Match      93.2%; Score 2111; DB 10; Length 2193;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2193; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

```

```

Qy      1 ATGATGTTTCTGCTTTTAAACATCTTCAAGCTCAGATGATTAACATGTTCTT 60
Db      1 ATGATGTTTCTGCTTTTAAACATCTTCAAGCTCAGATGATTAACATGTTCTT 60
Qy      61 CTACTTCATTCATGCTTCTGATGATGCAAGATTTTGAAGCTGACTGAAGGTACAG 120

```

Db 61 CTACTTCATTTTCATGCTCTGATCAATGTCAAAGATTTTGCACTGACTCAAGTAGCATG 120
Qy 121 ATCACTCCCTTCATGCGCAAAAGATATTTTCCCTGCGGAATCTTACCAAGTCTTACC 180
Db 121 ATCACTCCCTTCATGCGCAAAAGATATTTTCCCTGCGGAATCTTACCAAGTCTTACC 180
Qy 181 CGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAAACGGGGCGGACGAAAGAACTGT 240
Db 181 CGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAAACGGGGCGGACGAAAGAACTGT 240
Qy 241 GGTGCACTAGTGTGATGCGGACCATATTTGGGACACATGTCAGAAATGCTTACAGCTG 300
Db 241 GGTGCACTAGTGTGATGCGGACCATATTTGGGACACATGTCAGAAATGCTTACAGCTG 300
Qy 301 GCTTAAACACAGAGTCTTTCTTAAACAGTATCCACATGCTGTGACTGCAAAAGAACT 360
Db 301 GCTTAAACACAGAGTCTTTCTTAAACAGTATCCACATGCTGTGACTGCAAAAGAACT 360
Qy 361 GAATGGAATGTGTAATGTGATCTTAAAGTCTGTGCGGATGATTTCTTAAAGTGTGACA 420
Db 361 GAATGGAATGTGTAATGTGATCTTAAAGTCTGTGCGGATGATTTCTTAAAGTGTGACA 420
Qy 421 TTACTGCTCTTAAAGAAACAAATCCACAGTCTTCCAGATTAAGTTCATCAAAATAC 480
Db 421 TTACTGCTCTTAAAGAAACAAATCCACAGTCTTCCAGATTAAGTTCATCAAAATAC 480
Qy 481 ACAAACTTAAAGATATTTCTGACATATATGCAATTAAGACATATCCAGAAAGCA 540
Db 481 ACAAACTTAAAGATATTTCTGACATATATGCAATTAAGACATATCCAGAAAGCA 540
Qy 541 TTTTGTGATTAATGTAATCTGCAAAATATTAATCTCAACACAGTGTGATCAACACCTG 600
Db 541 TTTTGTGATTAATGTAATCTGCAAAATATTAATCTCAACACAGTGTGATCAACACCTG 600
Qy 601 AGAAGTGAATTTGAAAGAACTTACATGCTTAACTGCTTAACTTGAATGACATCA 660
Db 601 AGAAGTGAATTTGAAAGAACTTACATGCTTAACTGCTTAACTTGAATGACATCA 660
Qy 661 ATAAACAGAAATTTACAGGCTTGTGATGCGGATTAATCTCTGTTTCTGCTATG 720
Db 661 ATAAACAGAAATTTACAGGCTTGTGATGCGGATTAATCTCTGTTTCTGCTATG 720
Qy 721 GTTAATTAATCTTAAAGCTCTTCCACAGATGTGTGCCAAATGCTCAACTCAAC 780
Db 721 GTTAATTAATCTTAAAGCTCTTCCACAGATGTGTGCCAAATGCTCAACTCAAC 780
Qy 781 TGGGTGATTTGGAAGGCAATGAATTAAGTATCTCAAAATTTCAACGTTTCTGTGTC 840
Db 781 TGGGTGATTTGGAAGGCAATGAATTAAGTATCTCAAAATTTCAACGTTTCTGTGTC 840
Qy 841 GATGCTGACAGTGTCTGTTCTGCTGAATCAAAATGTTGTTGTCAGAGAAAGCA 900
Db 841 GATGCTGACAGTGTCTGTTCTGCTGAATCAAAATGTTGTTGTTGTCAGAGAAAGCA 900
Qy 901 TTTTCTTCACTTAAAAATTTAGAGACATGATCTGTCTAGCAATTCGATTAACGAGCTA 960
Db 901 TTTTCTTCACTTAAAAATTTAGAGACATGATCTGTCTAGCAATTCGATTAACGAGCTA 960
Qy 961 TCACCTCACCTTTTAAAGCTTGAAGCTTCTACAAAGCTGAACCTGTACATCAATCTCT 1020
Db 961 TCACCTCACCTTTTAAAGCTTGAAGCTTCTACAAAGCTGAACCTGTACATCAATCTCT 1020
Qy 1021 CTATGATATCTTCAAGAAACAGTGTGAAGCTTAAACAACTTCACTGTGAGCTG 1080
Db 1021 CTATGATATCTTCAAGAAACAGTGTGAAGCTTAAACAACTTCACTGTGAGCTG 1080
Qy 1081 GAAAGATGAGATTTCAAAATATTAACACAGATGTTTCAACGCTGAAAGAACTTTCT 1140
Db 1081 GAAAGATGAGATTTCAAAATATTAACACAGATGTTTCAACGCTGAAAGAACTTTCT 1140
Qy 1141 CACATTTATTTCAAAACCTTTGATCTGCTATGCTCCCATGTCGGAATATGATG 1200
Db 1141 CACATTTATTTCAAAACCTTTGATCTGCTATGCTCCCATGTCGGAATATGATG 1200
Qy 1201 GATTTATTTTCAAAACCTTTGATCTGCTATGCTCCCATGTCGGAATATGATG 1260
Db 1201 GATTTATTTTCAAAACCTTTGATCTGCTATGCTCCCATGTCGGAATATGATG 1260

Qy 1201 CCTTGAACGAGGCAATTTCTTCAATTGAGGACCTTTGCTTAACAAATATCTGAGATA 1260
Db 1201 CCTTGAACGAGGCAATTTCTTCAATTGAGGACCTTTGCTTAACAAATATCTGAGATA 1260
Qy 1261 TTTGTCTGGATTAATGCTTTCTTCAATTAATGCTTTGGAATCTTTTGTCTATGAGAGA 1320
Db 1261 TTTGTCTGGATTAATGCTTTCTTCAATTAATGCTTTGGAATCTTTTGTCTATGAGAGA 1320
Qy 1321 TCTTCAATTAAGCTGAAATTAACATCAAGCTATGTCATCAAAATCTTTGTTGTC 1380
Db 1321 TCTTCAATTAAGCTGAAATTAACATCAAGCTATGTCATCAAAATCTTTGTTGTC 1380
Qy 1381 GATTCCTGATGAGGATTTACTGTTCTTTGTTGCAATTTGATTAATAATACGAGGG 1440
Db 1381 GATTCCTGATGAGGATTTACTGTTCTTTGTTGCAATTTGATTAATAATACGAGGG 1440
Qy 1441 CAGTATCAGAAATATGCTTCTGCTGTGATGAGAGCGTGCAGCTGAGGCTTC 1500
Db 1441 CAGTATCAGAAATATGCTTCTGCTGTGATGAGAGCGTGCAGCTGAGGCTTC 1500
Qy 1501 CTGGCAGTGTGCTCAACGGAAGTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Db 1501 CTGGCAGTGTGCTCAACGGAAGTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Qy 1561 TTCTGTGCTATGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Db 1561 TTCTGTGCTATGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Qy 1621 ATCTCATTTGATCTGATGAGGAGATTTTAAATAGCTGTAATCTCAATTTGGAATAG 1680
Db 1621 ATCTCATTTGATCTGATGAGGAGATTTTAAATAGCTGTAATCTCAATTTGGAATAG 1680
Qy 1681 GATTAATTTGGAATCTTAAATGGAAGAAATGAGAGATGTTCCACTTATTAAGACCA 1740
Db 1681 GATTAATTTGGAATCTTAAATGGAAGAAATGAGAGATGTTCCACTTATTAAGACCA 1740
Qy 1741 ACAGAAATTTGGAAGCAAGGATTTCTTGAATTTCTAGATGTAATCTGCTG 1800
Db 1741 ACAGAAATTTGGAAGCAAGGATTTCTTGAATTTCTAGATGTAATCTGCTG 1800
Qy 1801 GCTTTCTCATCTGATGTTTCTGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1801 GCTTTCTCATCTGATGTTTCTGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Qy 1861 TTTTATTAATGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1920
Db 1861 TTTTATTAATGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1920
Qy 1921 TTTTATTAATGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1980
Db 1921 TTTTATTAATGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1980
Qy 1981 TTTTATTAATGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2040
Db 1981 TTTTATTAATGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2040
Qy 2041 CAGTAAACAGGCTTGAATCAATCTCTTAACTCAACCAACCTTTTAAAGAC 2100
Db 2041 CAGTAAACAGGCTTGAATCAATCTCTTAACTCAACCAACCTTTTAAAGAC 2100
Qy 2101 AAGTGAACAGCTGCTCAACCAATCAAGGAAATCAATTTTCAAAATTAAGAA 2160
Db 2101 AAGTGAACAGCTGCTCAACCAATCAAGGAAATCAATTTTCAAAATTAAGAA 2160
Qy 2161 AGTTATCAATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db 2161 AGTTATCAATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Qy 2221 AACAAATTAACCTTGAAGACATTAATGAACCAAGTTTCTAG 2285
Db 2221 AACAAATTAACCTTGAAGACATTAATGAACCAAGTTTCTAG 2285
Qy 2285 AACAAATTAACCTTGAAGACATTAATGAACCAAGTTTCTAG 2349
Db 2285 AACAAATTAACCTTGAAGACATTAATGAACCAAGTTTCTAG 2349

RESULT 6
US-09-965-536A-5
Sequence 5, Application US/0965536A
Publication No. US2003027323A1
GENERAL INFORMATION:
APPLICANT: FEDER, J. N.
APPLICANT: MINTER, G.
APPLICANT: RAMANATHAN, C. S.
APPLICANT: HAWKEN, D. R.
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGRPMY5,
TITLE OF INVENTION: EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
FILE REFERENCE: D0041NP
CURRENT APPLICATION NUMBER: US/09/965,536A
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/235,713
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/261,781
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/306,605
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/310,436
PRIOR FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 2142
TYPE: DNA
ORGANISM: Homo sapiens
US-09-965-536A-5

Query Match 90.9%; Score 2060; DB 11; Length 2142;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 52 ATGTTCTTCTACTTCAATTCATGCTTCTGATCAATGCAAGATTTTGGCACTGACAA 111
DB 1 ATGTTCTTCTACTTCAATTCATGCTTCTGATCAATGCAAGATTTTGGCACTGACAA 60
QY 112 GGTGACATGATCACTCTCTTATGCAAAAGATTTTCCGTGGGAATCTTACCAAG 171
DB 61 GGTGACATGATCACTCTCTTATGCAAAAGATTTTCCGTGGGAATCTTACCAAG 120
QY 172 TGTCTACCCGAGCTTTTCACTGATGCAAGATGACTGTGGAAAGCGGCGGAGAA 231
DB 121 TGTCTACCCGAGCTTTTCACTGATGCAAGATGACTGTGGAAAGCGGCGGAGAA 180
QY 232 GAGAACTGTGTGACACTAGTGTGGGCGCAATATTTGGCACTGCAATGCAATGCT 291
DB 181 GAGAACTGTGTGACACTAGTGTGGGCGCAATATTTGGCACTGCAATGCAATGCT 240
QY 292 AACAGCTGTGCTTAAACAGAGAGTCTTTCTTAAACAGATTCACAAATGCTGATGC 351
DB 241 AACAGCTGTGCTTAAACAGAGAGTCTTTCTTAAACAGATTCACAAATGCTGATGC 300
QY 352 AAAGAACTGAATGTGAATGTGAATGTGAATGTGAATGTGAATGTGAATGTGAAT 411
DB 301 AAAGAACTGAATGTGAATGTGAATGTGAATGTGAATGTGAATGTGAATGTGAAT 360
QY 412 AATGTGACATTAATGCTCTTAAAGAAACAAATTCACAGTCTTCCAGATTAAGTTTC 471
DB 361 AATGTGACATTAATGCTCTTAAAGAAACAAATTCACAGTCTTCCAGATTAAGTTTC 420
QY 472 ATCAATATCACAAAATCTTAAAGATTTTCTTCAAGATTAATGCAATGCAATGCAAT 531
DB 421 ATCAATATCACAAAATCTTAAAGATTTTCTTCAAGATTAATGCAATGCAATGCAAT 480
QY 532 AGGAAGCATTTTGTGATTAATGTAAATCTGCAATATTAATCTCAACCAACTGATC 591
DB 481 AGGAAGCATTTTGTGATTAATGTAAATCTGCAATATTAATCTGCAATATTAATCT 517
QY 592 AGAAGCTGAGACTGTGAATTTCAAAAGATTAATGCAATGCAATGCAATGCAATGCAAT 651

DB 518 -----TAATCTAGAT 528
QY 652 GACAACTCAATACAGAAATTTTCAAGGCTTGTATACGGATTAATTTCTTGTTC 711
DB 529 GACAACTCAATACAGAAATTTTCAAGGCTTGTATACGGATTAATTTCTTGTTC 588
QY 712 CTGTCTAATGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771
DB 589 CTGTCTAATGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 648
QY 772 CAATCAATGCTGTGATTTGAAGGCAATAGAAATTAATTAATTAATTAATTAATTAAT 831
DB 649 CAATCAATGCTGTGATTTGAAGGCAATAGAAATTAATTAATTAATTAATTAATTAAT 708
QY 832 CTGTCTGATTTGCTCACTGCTGCTTCTTCAAGAAATTAATTAATTAATTAATTAAT 891
DB 709 CTGTCTGATTTGCTCACTGCTGCTTCTTCAAGAAATTAATTAATTAATTAATTAAT 768
QY 892 GAGAAACATTTTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 951
DB 769 GAGAAACATTTTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 828
QY 952 ACGAGCTATCACTCACTTTTAAAGCTTGAAGCTTCAAAAGCTGAACCTGTCA 1011
DB 829 ACGAGCTATCACTCACTTTTAAAGCTTGAAGCTTCAAAAGCTGAACCTGTCA 888
QY 1012 TCCATCTCTTATGATTTTCAAGAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1071
DB 889 TCCATCTCTTATGATTTTCAAGAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 948
QY 1072 CTAGACCTGAAAGATTAAGATTTCAATATTAACACAGATTTTCAACCACTGAAG 1131
DB 949 CTAGACCTGAAAGATTAAGATTTCAATATTAACACAGATTTTCAACCACTGAAG 1008
QY 1132 AATCTTCTCACTTATTTCAAAAATTTGATGATCTGCTTCAAGCTTCAAGCTTCAAG 1191
DB 1009 AATCTTCTCACTTATTTCAAAAATTTGATGATCTGCTTCAAGCTTCAAGCTTCAAG 1068
QY 1192 AATGATGATGCTTGAAGGCAATTTCTTCAATTAAGAGACCTTGGCTTAACATATC 1251
DB 1069 AATGATGATGCTTGAAGGCAATTTCTTCAATTAAGAGACCTTGGCTTAACATATC 1128
QY 1252 CTCAGAAATTTGCTGAGTTATAGCTTTCATTAAGCTTGTGAAGCTTGTGATTC 1311
DB 1129 CTCAGAAATTTGCTGAGTTATAGCTTTCATTAAGCTTGTGAAGCTTGTGATTC 1188
QY 1312 GGCATGATGCTTCTTCAATTAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 1371
DB 1189 GGCATGATGCTTCTTCAATTAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 1248
QY 1372 TGTGATGCTGATGCTGATGAGGCTTGAAGCTTGTGATGAGCTTGTGATGAGCTTGTGAT 1431
DB 1249 TGTGATGCTGATGCTGATGAGGCTTGAAGCTTGTGATGAGCTTGTGATGAGCTTGTGAT 1308
QY 1432 TACCGAGGCAATGATGAGATGAGCTTGTGATGAGGCAAGCTGAGTCCGCTTC 1491
DB 1309 TACCGAGGCAATGATGAGATGAGCTTGTGATGAGGCAAGCTGAGTCCGCTTC 1368
QY 1492 ATGAGGCTTCTGAGCTTGTGATGAGGCAAGCTTGTGATGAGGCAAGCTTGTGATGAGCT 1551
DB 1369 ATGAGGCTTCTGAGCTTGTGATGAGGCAAGCTTGTGATGAGGCAAGCTTGTGATGAGCT 1428
QY 1552 TTGGAAGATTTCTGATGATGCTTCCCTTCAATTAATGCACTGAGAAACGGCAG 1611
DB 1429 TTGGAAGATTTCTGATGATGCTTCCCTTCAATTAATGCACTGAGAAACGGCAG 1488
QY 1612 ACTTCAATGCTTCAATTTGATCTGATGAGGCGGATTTTAAATGCTGATTAATTCATTT 1671
DB 1489 ACTTCAATGCTTCAATTTGATCTGATGAGGCGGATTTTAAATGCTGATTAATTCATTT 1548
QY 1672 TGAATTAAGATTAATTTGGAACCTTTATGGAAGAAATGAGATAGTTTCCACTTAT 1731
DB 1549 TGAATTAAGATTAATTTGGAACCTTTATGGAAGAAATGAGATAGTTTCCACTTAT 1608

1141 CAGATTATTTTCAAAAACCTTGCACTAGTCTCTATGCTCCCATGTCGGAATATGTATG 1200
Db 925 CAGATTATTTTCAAAAACCTTGCACTAGTCTCTATGCTCCCATGTCGGAATATGTATG 984
Qy 1201 CCCTTGACGACGCGCATTTCTTCACTTGAAGACCTCTTGCGTAACAATATCTCAGAAATA 1260
Db 985 CCCTTGACGACGCGCATTTCTTCACTTGAAGACCTCTTGCGTAACAATATCTCAGAAATA 1044
Qy 1261 TTTGTCTGGGTATAGCTTTCATTTACCTGCTTTGAAAATCTTTTGTGATGGCATGAGA 1320
Db 1045 TTTGTCTGGGTATAGCTTTCATTTACCTGCTTTGAAAATCTTTTGTGATGGCATGAGA 1104
Qy 1321 TCTTTATTTAAGCTGAAAATACAACTCAGCGTATGCGCATCAAAAACCTTTGTGGCT 1380
Db 1105 TCTTTATTTAAGCTGAAAATACAACTCAGCGTATGCGCATCAAAAACCTTTGTGGCT 1164
Qy 1381 GATTGCTGATGGGTGTTTACTTGTCTTGTGGCATTTTGCATATATAACCGAGGG 1440
Db 1165 GATTGCTGATGGGTGTTTACTTGTCTTGTGGCATTTTGCATATATAACCGAGGG 1224
Qy 1441 CAGATCAGAAGTATGCTTGTCTGTGATGAGAGCGTGACGTGCGCTCATGGGGTTC 1500
Db 1225 CAGATCAGAAGTATGCTTGTCTGTGATGAGAGCGTGACGTGCGCTCATGGGGTTC 1284
Qy 1501 CTGGCCATGCTGTCCACCGAAGTCTCTTCTGTACTGACCTTACTGACTTTGGAGAG 1560
Db 1285 CTGGCCATGCTGTCCACCGAAGTCTCTTCTGTACTGACCTTACTGACTTTGGAGAG 1344
Qy 1561 TTTCTGTGATATGCTTCCCTTCAGTAACTTGAAGTATGTTTCCCATTTTATGACCAA 1620
Db 1345 TTTCTGTGATATGCTTCCCTTCAGTAACTTGAAGTATGTTTCCCATTTTATGACCAA 1404
Qy 1621 ATCTCATTTGCACTGATGAGCGGGAATTTTAACTGTATTCATTTTGGAAATAG 1680
Db 1405 ATCTCATTTGCACTGATGAGCGGGAATTTTAACTGTATTCATTTTGGAAATAG 1464
Qy 1681 GATTATTTTGGAACTTTTATGGAAGAAATGAGTATGTTTCCCATTTTATGACCAA 1740
Db 1465 GATTATTTTGGAACTTTTATGGAAGAAATGAGTATGTTTCCCATTTTATGACCAA 1524
Qy 1741 ACAGAGAATATTTGAGCAAGGATATCTCTTGGAAATTTTCTAGGTGTGAACCTTGTG 1800
Db 1525 ACAGAGAATATTTGAGCAAGGATATCTCTTGGAAATTTTCTAGGTGTGAACCTTGTG 1584
Qy 1801 GCTTTTCTCATATGTTGTTTCTCTATATCTATGTTCTGTCTCATTTCAAAAACCGCC 1860
Db 1585 GCTTTTCTCATATGTTGTTTCTCTATATCTATGTTCTGTCTCATTTCAAAAACCGCC 1644
Qy 1861 TTGCAGACCAAGAAATGAGAAATGTTTGGAAAGAGGTGGCTGTGCAAAATCGTTTC 1920
Db 1645 TTGCAGACCAAGAAATGAGAAATGTTTGGAAAGAGGTGGCTGTGCAAAATCGTTTC 1704
Qy 1921 TTTTATTAAGTGTCTCTGATGCAATCTGCTGATTCCTGTATTTAGTTAAATCCTT 1980
Db 1705 TTTTATTAAGTGTCTCTGATGCAATCTGCTGATTCCTGTATTTAGTTAAATCCTT 1764
Qy 1981 TCCCTCTTCCGGGTGGAATACAGACACATGACTTCTGTGATGATTTTCTTCTT 2040
Db 1765 TCCCTCTTCCGGGTGGAATACAGACACATGACTTCTGTGATGATTTTCTTCTT 1824
Qy 2041 CCAGTTAACAGTGTGATTCATCTCTATATCTCTCAACCAATTTTAAAGAC 2100
Db 1825 CCAGTTAACAGTGTGATTCATCTCTATATCTCTCAACCAATTTTAAAGAC 1884
Qy 2101 AAGTTGAAAACGCTGTCACAAACATCAGAGAAATCAATTTTCAAAATTTAAAAAAA 2160
Db 1885 AAGTTGAAAACGCTGTCACAAACATCAGAGAAATCAATTTTCAAAATTTAAAAAAA 1944
Qy 2161 AGTTATATCATCATGATGTGTGATAGAGACTCTTCTGCGAAACCTTGGGGTTTG 2220
Db 1945 AGTTATATCATCATGATGTGTGATAGAGACTCTTCTGCGAAACCTTGGGGTTTG 2004

Qy 2221 AACAAATATACCTTGAGACAGTATATGAAACCAAGTTCTAG 2265
Db 2005 AACAAATATACCTTGAGACAGTATATGAAACCAAGTTCTAG 2049
RESULT 8
US-10-229-735-4
; Sequence 4, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
; APPLICANT: Baylor College of Medicine
; APPLICANT: Agoulnik, Alexander I.
; TITLE OF INVENTION: The GREAT Gene and Protein
; FILE REFERENCE: 7572/73263
; CURRENT APPLICATION NUMBER: US/10/229,735
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/315,696
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/351,432
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2539
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-229-735-4

Query Match 69.7%; Score 1579.8; DB 14; Length 2539;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1831; Conservative 0; Mismatches 397; Indels 3; Gaps 1;

Qy 32 TCAGCTCAGATGATTAAGTCTTCTTCACTTCAATTCATTCGCTTGATGATGCA 91
Db 174 TCAGGCGCACACCGCGCGAATGCGCTCTTCACTTCAATTCCTTTCAGAGAGTCA 233
Qy 92 AAGATTTTGCATGACTGACTCAAGTAGCATGATCACTCTTCATGCAAAAAGATATTTTC 151
Db 234 AAGATTTTGCATGACTGACTGACAGCATATGATGCTCTCTGATGCCCAAGGATATTTTC 293
Qy 152 CTTGTGGAAATCTTACCAAGTCTTACCCGAGCTTTTCACTGTATGCGAAGATGACT 211
Db 294 CTTGTGGAAATCTTACCAAGTCTTACCCGAGCTTTTCACTGTATGCGAAGATGACT 353
Qy 212 GTGGGAAACGGGGGCGAGCAAGGAACGTGTGTGACACTAGTGAATGGGCGACCATATTTG 271
Db 354 GCGGGAATGCTGCGACGAGGACACTGTGTGTGACACTAGTGAATGGAACCATATTTG 413
Qy 272 GCAAGTGCATGAAATGCTAACAGCGTGGCTTTAACACAGAGTCTTTCTAAACAGT 331
Db 414 GCAAGTGCATGAAATGCTAACAGCGTGGCTTTAACACAGAGTCTTTCTAAACAGT 473
Qy 332 ATCCACATGCTGTGACTGCAAGAAACCTGAATTTGAAATGTGAAATGTGACTTAAAT 391
Db 474 ATCCACATGCTGTGACTGCAAGAAACCTGAATTTGAAATGTGAAATGTGACTTAAAG 533
Qy 392 CTGTGCGAATGATTTTCAACATGTGACATTTACTGTCTCTTAAGAAAAACAAATCCACA 451
Db 534 CTGTGCGAATGATTTTCAACATGTGACATTTACTGTCTCTTAAGAAAAACAAATCCACA 593
Qy 452 GTCTTCAGATTAAGTTTTCATCAATTAACAAACCTTAAAGAAATATTTCTTCAAGATA 511
Db 594 GACTTCAGTCAAGGCTTTCAGAGATTAACAAACCTTAAAGAAATATTTCTTCAAGATA 653
Qy 512 ATTGCATTAACATATCCAGAAAGCAATTTTGTGATTAATGTAATCTGCAAAATATAT 571
Db 654 ACTGCATCAACATCTCCAGAGAGCAATTTCTGTGATTAACATTAATCTAACAATACTGT 713
Qy 572 ATCTCAACCAACATGATCAACACCTCAGACCTGGAATATTCAAAGACTTACATCAGC 631
Db 714 ATCTCAACCAACATGATTAATCTCTCAGGCTGGAATATTCAAAGACTTACATCAGC 773
Qy 632 TAACTTGGCTAATTTAGATGACATCAATTAACAGAAATTTCAACAGGCTTGTATCGG 691


```

Db      774  TTGCTGGCTAATTTAGATGACACCGATCACCAGATCTCAGAGAGCTTTATAG 833
Qy      692  GATTAATTCCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 751
Db      834  GGTTAATTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 893
Qy      752  AGATGTGTCCCAATGCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 811
Db      894  CATGTGTGTCAATGCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 953
Qy      812  ATCTCAAAATTTCAAGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 871
Db      954  ACATTAAGAACTCCACCTTCTTAACGTGCGACTGCGCTGCGCTGCGCTGCG 1013
Qy      872  ATCAAAATTTGTTTGTTCAGAGAAACATTTTCTTCACTTAATAAAATTTAG 931
Db      1014  ATCAAAATTTGTTTGTTCAGAGAAACATTTTCTTCACTTAATAAAATTTAG 1073
Qy      932  ATCTGTAGCAATACGATAACGAGCTATCACTCACTTTTAAAGCTTGAAGCTTC 991
Db      1074  ACCTGTAGCAATATGATAACAAACCTCCAGTCCACTTTTCAAGCACTTGCATCTTC 1133
Qy      992  TACAAAAGCTGAACCTGTATCCAAATCTCTTTATGTATCTTCAAGAAACAGTTG 1051
Db      1134  TCCGAAAGCTGAACCTGTATCCAAACCTCTTGTATGTATCCAAAGAAACAGTTG 1193
Qy      1052  GTCTTAAACCACTTCACTGTCTGAGACCTGAAAGGATGAGATTCGAAATTAAC 1111
Db      1194  GTCTCAAAACCACTTCACTGTCTGAGACCTGAAAGGATGAGATTCGAAATTAAC 1253
Qy      1112  GAATGTTTCAACCCATGAGAAATCTTTCTCACTTATTTCAAAAACCTTGCATCT 1171
Db      1254  GAATGTTTCAACCCATGAGAAACCTTTCTCACTTATTTGAAAACCTTGCATCT 1313
Qy      1172  CCTATGCTCCCAATGCTGGAATATGTATGCTTGAAGGAGCACTTTCTTCACTT 1231
Db      1314  CCTATGCTCCCAATGCTGGAATATGTATGCTTGAAGGAGCACTTCTTCTGCTG 1373
Qy      1232  ACCCTTGGCTAACAAATATCTCCAGAAATTTGTGTGGGTTATAGCTTCAATCT 1291
Db      1374  ACCCTTGGCTAACAAATATCTCCAGAAATTTGTGTGGGTTATAGCTTCAATCT 1433
Qy      1292  TTGGAATCTTTTGTCTATGAGATGAGATCTTTTCACTTAAAGCTGAAATACA 1351
Db      1434  TTGGAATCTTTTGTCTATGAGATGAGATCTTTTCACTTAAAGCTGAAATACA 1493
Qy      1352  CTATGTCCATCAAAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1411
Db      1494  CTATGTCCATCAAAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1553
Qy      1412  TTGGAATCTTTTGTCTATGAGATGAGATCTTTTCACTTAAAGCTGAAATACA 1471
Db      1554  TTGGAATCTTTTGTCTATGAGATGAGATCTTTTCACTTAAAGCTGAAATACA 1613
Qy      1472  AGAGGTGCAAGTGCCTCATGAGGTTCTTGCCAGTCTGTCAACGAGTCTGTCT 1531
Db      1614  AGAGGTGCAAGTGCCTCATGAGGTTCTTGCCAGTCTGTCAACGAGTCTGTCT 1673
Qy      1532  TGCTAATGACCTTCTTGAAGTTTGTGAGAGTTTCTGTGTCTTCTTCTTCTTCA 1591
Db      1674  TGCTAATGACCTTCTTGAAGTTTGTGAGAGTTTCTGTGTCTTCTTCTTCTTCA 1733
Qy      1592  TTGCACTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1651
Db      1734  TTGCACTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1793
Qy      1652  TAAATGCTTAATTCATTTTGAATTAAGATTAATTTTGAATTAATTTTGAATTA 1711
Db      1794  TCAATGCGGCGGCTTCTGTTCAACGAGAGATTAATTTGCGCACTTTATGGA 1853
Qy      1712  GAGTATGTTTCCATTTATATGACCAACGAGAGATTTGGAAGCAAGGTATTTCT 1771

```

```

Db      1854  GAGTGTGCTTCCACTTATTAATGACCAAGAGATTTTGAAGTGAAGGATCTCC 1913
Qy      1772  TTGGAATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1831
Db      1914  TTGGAATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1973
Qy      1832  CTATGTCTTCTTCAATCAAAACCGCTTGAAGTGAAGTGAAGTGAAGTGAAG 1891
Db      1974  CAAATGTTCTCTCATTCATTAACAGCCCTTGAAGTGAAGTGAAGTGAAGTGA 2033
Qy      1892  GAAGAGAGTGTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1951
Db      2034  GAAGAGAGTGTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2093
Qy      1952  GAAATTCCTGATTTGATTAATAATCTTCCCTTCCGAGGAGGAGGAGGAGGAG 2011
Db      2094  GAAATTCCTGATTTGATTAATAATCTTCCCTTCCGAGGAGGAGGAGGAGGAG 2153
Qy      2012  TGAATTCCTGATTTGATTAATAATCTTCCCTTCCGAGGAGGAGGAGGAGGAG 2071
Db      2154  TGAATTCCTGATTTGATTAATAATCTTCCCTTCCGAGGAGGAGGAGGAGGAG 2213
Qy      2072  ATATCTCAACCACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2131
Db      2214  ATATCTCAACCACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2273
Qy      2132  GAAATCAATTTTCAAAATTAATAATAATTTTGAAGTGAAGTGAAGTGAAGTGA 2191
Db      2274  GAAATCAATTTTCAAAATTAATAATAATTTTGAAGTGAAGTGAAGTGAAGTGA 2333
Qy      2192  ACTTCTTCTTCAAACTTGGGTTTGAACAAATTAACCTTGAAGTGAAGTGAAG 2251
Db      2334  A--GTCTCACTTAACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2390
Qy      2252  AACGAGTTTC 2262
Db      2391  AGCGGTCTCC 2401

```

RESULT 9

```

US-09-928-175-19
; Sequence 19, Application US/09928175
; Patent No. US2002123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ..(2211)
; NAME/KEY: sig_peptide
; LOCATION: (1) ..(57)

```

Query Match 69.7%; Score 1579; DB 10; Length 2214;

Best Local Similarity 82.5%; Pred. No. 0;

Matches 1823; Conservative 0; Mismatches 355; Indels 3; Gaps 1;

Qy 52 ATGTTCTTTCTTCAATTTCAATCTTCTGATCAATGATCAAAAGATTTTGCACGTACACAA 111

Db	1	ATGGGCTCTCACTTCATGTCATCTCTTGACAAGGTCCTCAAGATTTTGGCACTGGCTGAC	60
Qy	112	GGTAGCATGATCACTCTTCATGTCCTCAAAAGATTTTCCCTGGGAACTTACCAAG	171
Db	61	AGCAGTATGATGGCTCTCTGTGGCCCCAAGGATTTTCCCTGTGGGAATCTCACAAA	120
Qy	172	TGCTTATCCCCGAGCTTTCACTGTGATGGCAAGATGACTGTGGGAACGGGGCGAGCGAA	231
Db	121	TGCTTGGCCCCGAGCTTTCACTGCACATGTGTGGATGATTCGGGAAATGGTGGCCGACAG	180
Qy	232	GAGAACTGGGAGCACTAGTGGATGGGCGACATATTTGGCAGCTGATGGAATGCT	291
Db	181	GACAACTGTGGTGAACATGATGGATGGAACACCAATTTTGGCAGTCCATGGGAATGTC	240
Qy	292	AACAGCGTGGCCTTAAACAAGGATGCTTTCTTAAACGATATCCAAATGCTGTGACTGC	351
Db	241	AATAAAGTGAACATTGACACAGAGAGTGTTCCTCAGCCAGTATCCACACACTGTTAATGTC	300
Qy	352	AAAGAACTGAAATTTGGAATGTGTAAATGATGACTTAAAGTCTGTGGCCGATATTTTCAAC	411
Db	301	AGAGAAATGAACTGAAATGTGTAAAGCTGACTTAAAGCTGTGGCCAAAGTGTTCACGC	360
Qy	412	AATGACATTTACTGCTCTTAAAGAAAAACAATCCAGTCTTCCAGATTAAGTTTC	471
Db	361	AACGTAACTTATCTCTCTTAAAGAAAAACAATACTCAAGCTTCCAGTCAAGTCTTC	420
Qy	472	ATCAATATCACAAAACCTTAAAAAGATTTCTTCAGCAATAATGCAATTAAGACATATCC	531
Db	421	AGCAGATACACAGAACTGAAAGATTAATCTTCAGCAACAATGATCACACATCTCC	480
Qy	532	AGGAAGAATTTTGGATTAATGTATCTGCAAAATTAATATCTCAACACAACTGCATC	591
Db	481	AGGAGAGATTCCTGTGGATACATATCTCAAAATCTGTATCTCAGGCAATACTGATTT	540
Qy	592	ACAACTCTCAGACTGGAAATATTCAAGAATTCAATCACTGACTAATCTGCTATGAT	651
Db	541	ACCTCTCAGGCTGGGATATTCAAAAGCTGCATCAGCTGTGCTGAATTTTATGAT	600
Qy	652	GACATATCCAAATPACAGAAATTTCAAGGCTGTTTAAAGGATTAATTCCTGTTTTTC	711
Db	601	GACAACTCGATACCAAGATCTCAAGAAATCTTTATGGGTTTAACTCTCTGTTTTTC	660
Qy	712	CTGTATGTTAATTAATTAATTAATTAAGACTCTTCCCAAGATGTGTGCCAAATGCT	771
Db	661	TGTGCAATGTTGGTAACTGACGAGCTCGAGGCCCTTCTGAAACATGTGTGCTCAGATGCT	720
Qy	772	CAATCCAACTGGGTGGATTTGGAAGCAATAGAAATTAAGATCTCAAAATTTACGTTT	831
Db	721	CAATCCAACTGGGTGGATCTGGCAAAACATGAAATTAAGATCAATPACAACTCCACTTC	780
Qy	832	CTGTGTGCGAATCCCTCAAGTGTGTTTTGCTCTAAGAAATCAAAATGTTTTGTTTCA	891
Db	781	CTAAGTGGACTCGCTCAAGGTTCTGTTTTCTGCTTAAGAAATCAAAATGTTTTGTTTCA	840
Qy	892	GAGAAACATTTTCTTCATTAATAATTTAGAGAACTGATCTGTCTAGCAATACGATA	951
Db	841	GAGAAACATTTTCTTCATTAATAATTTAGAGAACTGATCTGTCTAGCAATATGATA	900
Qy	952	ACGAGCTATCACTCACTTTTAAAGCTTGAAGCTTCAACAAAACCTGAACCTGCA	1011
Db	901	ACAAAATCTCCAGTCCACTTTTTCAGGACCTTCATTTCTTCAGAACTGAACCTGTCA	960
Qy	1012	TCCAATCTCTTATGATCTTCAACAAGAACGATTTGAAGCTTTAAACAACTTCAGTCT	1071
Db	961	TCCAATCTCTCTTATGATCTTCAACAAGAACGATTTGAAGCTTTAAACAACTTCAGTCT	1020
Qy	1072	CTAGACCTGGAAAGATGAGATTCCAATATAAACAACGAATGTTTCAACCCATGAG	1131
Db	1021	CTAGACCTGGAAAGATGAGATTCCAATATAAACAACGAATGTTTCAACCCATGAG	1080
Qy	1132	AATCTTCTCAATTAATTTCAAAACCTTGGATACTGCTCTATGCTCCCATGTCCGA	1191

Db	1081	AACCTTCTCAGATTATTTGAAAACTTTGATAGTCTGCTCTATGTCCTCCCATGTCGA	1140
OY	1192	ATATGTAATGCCCTTGAACGCAACGGCATTTCTCATTTTGAAGACCTTGGCTAACAAATATC	1251
Db	1141	ATCTGTATGCCGTCGACTGATAGTATTTCTTTCGTCTGAGGACCTCTTGGCTAACGGTATC	1200
OY	1252	CTCGAAATTTGTGTCGGGTATATAGCTTTCAATACCTGCTTTGGAAATCTTTTGTCAAT	1311
Db	1201	CTCAGAGTCTGTCTGGGTATATAGCTTTCATTAACCTGCGGTGGAAATTTCTTGTGATA	1260
OY	1312	GGCATGAGATCTTTCATTTAAAGCTGAAATACAACCTACGCTATGTCATCAAAATCCTT	1371
Db	1261	GCCGTGAGATCTCTCATTTAAGCTGAGAAATACACTACGCTATGTCATCAAAATCCTT	1320
OY	1372	TGTTGTGCTGATTCCTCGATAGGTGTTTACTGTTCTTTGTTGGCATTTTCGATATMAAA	1431
Db	1321	TGTTGTGCGCATTCCTCGATAGGGGTGTACTGTTCTTCGGTGGCGCTTGTGATCAAG	1380
OY	1432	TACCGAGGGCATATACAGAGTATAGCCTGTGCTGATAGAGAGACGTCAGTCGGCTCTC	1491
Db	1381	TACCGAGGGCATATACAGAGTATAGCCTGTGCTGATAGAGAGTGTGCTTCGCGCTTG	1440
OY	1492	ATGGGGTTTCTCGGCGCATGCTGTCCACCGAAATCTGTCTTCTACTGACCTTACTGACT	1551
Db	1441	CTGGGCTTCTCGGCGCATGCTGTCCACAGAGTCTCGGTGCTGTCTGACATTCCTGACG	1500
OY	1552	TTGAGAAAGTTCTCTGTGCTATGTCTTCCCTTCACTAACAATTGCACTGTGAAAAACGCGAC	1611
Db	1501	CTGAGAAAGTTCTCTGTGCTATGTCTTCCCTTCACTAACAACCTGTGCGGCAACGCGCAG	1560
OY	1612	ACCTCAGATCATCTCATTTGTCATCTGATGAGCGGGAATTTTAAATAGCTGTAAATCCATTT	1671
Db	1561	ACTGCTGTGGCCCTCGCAGACATCTGGGTGTGGGATTTCTCANAGGGCCGTTCCTGCTC	1620
OY	1672	TGGAATAGATTAATTTTGGAACTTTTATGCGAAAAATGAGATGTGTTTCCACTTAT	1731
Db	1621	ACCGAGAGGATTAATTTTTCGGCACTTTTATGCGAAAAATGAGATGTCTCTTCCCACTTCAT	1680
OY	1732	TATGACCAAAACAGAAATATATGAAAGCAAAAGGATATCTCTTGGAAATTTTCTTAGTGTCG	1791
Db	1681	TATGACCAAGCAAGATTTTGGAAATGAGAGTATCTCCTTGGGATTTTCTTAGTGTCG	1740
OY	1792	AACCTGCTGCTTTTCTCATATGTGTGTTTCTTATATCTATGTCTGTTCATTCAA	1851
Db	1741	AACCTGCTGCTTTTCTCTGATATGTATATTTCTATGTATCAACATGTCTGTCTCATTCAT	1800
OY	1852	AAAACCGCTTGCAGACCAACAGAGTAAAGGAATGTTTGGAAAGAGGTGGCTGTGCA	1911
Db	1801	AAAACCGCTTGCAGATCTGCAAGATGAGAGCCACATCCGGAAGAGAGTGGCTGTGCA	1860
OY	1912	AATCGTTTCTTTTATATAGTGTCTCTGATGCCATCTGTGAGATTCTGTATTTGTAGTT	1971
Db	1861	AAACGGTTCTTTTATATGATGTCTCTGATGCCATCTGTGAGATTCTGTATTTGTAGTT	1920
OY	1972	AAAATCTCTTCCCTCTCTCCGGGTGGAAATACCAACACATATGATTTCTGATATGTATTT	2031
Db	1921	AAAGATCTCTCTCTCTCAAGTGGAAATACCAAGCAACATCACTCTGTGATGTGATTT	1980
OY	2032	TTTTTCTCTCAGTTAACAGTGTCTTGAATCCATCTCTTATCTCTCACAAACCACTTT	2091
Db	1981	TTTTTCTCTCCGGTGAACAGGCTTTAAACCCCATCTCTTACATCTGAGGACCTCGTTT	2040
OY	2092	TTTAAAGCAAGTTGAAAACAGCTGCTGCACAACATACAGAGAAATCAATTTTCAAAAT	2151
Db	2041	TTTAAAGCAAGTTGAAAACAGTTGCTCACAACATGAGAGAAACCCATCTTCAAAAGTT	2100
OY	2152	AAAAAAAAGTTTATCTACATCATTTGTGTGATAGAGACTCTCTTCCCTGAAACTT	2211
Db	2101	AAGAGAAATTTTATCCGCTTCCATTTGTGTGTGACAGACA--GTCCTACCTTAAACTT	2157
OY	2212	GGGGTTTGAACAAATTAACCTTGGAGACGATTAATGAAACCAAGTTTTC	2262
Db	2158	GGAGTGTGAACAAATATGACCTTGGGAGACGATTAATGAGCCGCTCTCC	2208

RESULT 10
US-09-928-175-4
; Sequence 4, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1149)
US-09-928-175-4

Query Match
Blast Local Similarity 100.0%; Score 1149; DB 10; Length 1149;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 CAAGTAGCATGATCACTCTTCATGCCAAAAAGATATTTTCCCTGTGGATCTTACC 168
DB 1 CAAGTAGCATGATCACTCTTCATGCCAAAAAGATATTTTCCCTGTGGATCTTACC 60
QY 169 AAGTGTTCACCCGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAAACGGGCGGAC 228
DB 61 AAGTGTTCACCCGAGCTTTTCACTGTGATGCGAAGATGACTGTGGGAAACGGGCGGAC 120
QY 229 GAAGGAATGCTGTGACACTAGTGTGGAGGAGCACTATTTTGGCACTATGATGAAT 288
DB 121 GAAGGAATGCTGTGACACTAGTGTGGAGGAGCACTATTTTGGCACTATGATGAAT 180
QY 289 GCTAACAGGTGGCTTAAACAGAGAGTCTTTCTTAAACAGATGCAATGCTGTGAC 348
DB 181 GCTAACAGGTGGCTTAAACAGAGAGTCTTTCTTAAACAGATGCAATGCTGTGAC 240
QY 349 TGCAGAAAGAACTGATTTGAATGTGTAATGTGACTTAAAGTCTGTGCGATGATTTCT 408
DB 241 TGCAGAAAGAACTGATTTGAATGTGTAATGTGACTTAAAGTCTGTGCGATGATTTCT 300
QY 409 AACAAATGTACATTTACTGCTCTTAAAGAAAACAAATCCACAGTCTTCCAGATTAAGTT 468
DB 301 AACAAATGTACATTTACTGCTCTTAAAGAAAACAAATCCACAGTCTTCCAGATTAAGTT 360
QY 469 TTCATCAATATACAAACAACTTAAAGATATTTCTTACAGATTAATGATGACACATA 528
DB 361 TTCATCAATATACAAACAACTTAAAGATATTTCTTACAGATTAATGATGACACATA 420
QY 529 TCCAGAAAGCAATTTTGTGATTAATCTGCAAAATATTAATCTCAACCAACATGTC 588
DB 421 TCCAGAAAGCAATTTTGTGATTAATCTGCAAAATATTAATCTCAACCAACATGTC 480
QY 589 ATCAACACCTGAGACCTGGAATATCAAGACTTACACAGCTTACCTGGGCAATTTCTA 648
DB 481 ATCAACACCTGAGACCTGGAATATCAAGACTTACACAGCTTACCTGGGCAATTTCTA 540
QY 649 GATGACATTCATTAACAGAAATTTTCAACGCGCTTGTATTCAGGATTAATTCCTGTTT 708
DB 541 GATGACATTCATTAACAGAAATTTTCAACGCGCTTGTATTCAGGATTAATTCCTGTTT 600

QY 709 TTCTGTCTATGTTAATACTTAAGACCTTTCCCAAGCAGATGTGTCCCAATG 768
DB 601 TTCTGTCTATGTTAATACTTAAGACCTTTCCCAAGCAGATGTGTCCCAATG 660
QY 769 CCTCAACTCACTGGGAGATTTGGAAGGCAATGAATTAAGTATTCACAAATCTCAG 828
DB 661 CCTCAACTCACTGGGAGATTTGGAAGGCAATGAATTAAGTATTCACAAATCTCAG 720
QY 829 TTCTGTCTATGTTAATACTTAAGACCTTTCCCAAGCAGATGTGTCCCAATG 888
DB 721 TTCTGTCTATGTTAATACTTAAGACCTTTCCCAAGCAGATGTGTCCCAATG 780
QY 889 CCAAGAAAGACATTTTCTTCAATTAATAATTTAGAGACTGATGTCTTACGATTCG 948
DB 781 CCAAGAAAGACATTTTCTTCAATTAATAATTTAGAGACTGATGTCTTACGATTCG 840
QY 949 ATTAAGGAGCTATACCTTCACTTTTAAAGACTTGAAGCTTCAAAAGCTGACCTG 1008
DB 841 ATTAAGGAGCTATACCTTCACTTTTAAAGACTTGAAGCTTCAAAAGCTGACCTG 900
QY 1009 TCATCAATCTCTTATGATCTTCAACAAAGCAGTTGAAAGCTTAAACAACTTCAG 1068
DB 901 TCATCAATCTCTTATGATCTTCAACAAAGCAGTTGAAAGCTTAAACAACTTCAG 960
QY 1069 TCTCTAGCTCTGAAAGATGAGATTCGAATTAATAACACAGATTTTCAACCCATG 1128
DB 961 TCTCTAGCTCTGAAAGATGAGATTCGAATTAATAACACAGATTTTCAACCCATG 1020
QY 1129 AAGATCTTCTCACTTATTAATTAATAAACTTGTGATCTCTATGCTCCCATGTC 1188
DB 1021 AAGATCTTCTCACTTATTAATTAATAAACTTGTGATCTCTATGCTCCCATGTC 1080
QY 1189 CGAATATGATGCTGCTTACAGGAGCACTTTCTTCAATTTAGAGACTTGTGCTTAACAT 1248
DB 1081 CGAATATGATGCTGCTTACAGGAGCACTTTCTTCAATTTAGAGACTTGTGCTTAACAT 1140
QY 1249 ATCTTCAGA 1257
DB 1141 ATCTTCAGA 1149

RESULT 11
US-10-321-807-17
; Sequence 17, Application US/10321807
; Publication No. US20030166148A1
; GENERAL INFORMATION:
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; TITLE OF INVENTION: No. US20030166148A1-Endogenous, Constitutively Activated Human G
; FILE REFERENCE: AREN0086
; CURRENT APPLICATION NUMBER: US/10/321,807
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US/09/714,008
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,088
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,099
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,369
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/171,902
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,901
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,900
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/181,749

```
; PRIOR FILING DATE: 2000-02-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-321-807-17

Query Match      47.1%; Score 1066.4; DB 12; Length 1068;
Best Local Similarity 99.9%; Pred. No. 1.2e-270;
Matches 1067; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1198 ATGCTTGTGACGAGCGCATTTCTTCAATTTGAGAGACCTCTGGCTAACATATCTCACA 1257
DB 1 ATGCTTGTGACGAGCGCATTTCTTCAATTTGAGAGACCTCTGGCTAACATATCTCACA 60
QY 1258 ATATTTGTCTGGGTTATAGCTTTCATTAACCTGCTTGGAAATCTTTTGTGATTTGGCAT 1317
DB 61 ATATTTGTCTGGGTTATAGCTTTCATTAACCTGCTTGGAAATCTTTTGTGATTTGGCAT 120
QY 1318 AGATCTTTCATTAAGCTGAAATATACAACTCAGCTATGTCATCAAAATCTTTGTGT 1377
DB 121 AGATCTTTCATTAAGCTGAAATATACAACTCAGCTATGTCATCAAAATCTTTGTGT 180
QY 1378 GCTGATTTGCTGATGGGTGTTTACTGTCTTTGTTGGCATTTTGCATATTAATACCA 1437
DB 181 GCTGATTTGCTGATGGGTGTTTACTGTCTTTGTTGGCATTTTGCATATTAATACCA 240
QY 1438 GGGCAGATCAGAGATGAGCTGCTGCTGCTGAGAGAGCGTGAGAGCGGCTCATGGGG 1497
DB 241 GGGCAGATCAGAGATGAGCTGCTGCTGCTGAGAGAGCGTGAGAGCGGCTCATGGGG 300
QY 1498 TTCCTGGCCATGCTGCTCACAAGTCTCTGTTCTGCTACTGACCTTACCTTTGGAG 1557
DB 301 TTCCTGGCCATGCTGCTCACAAGTCTCTGTTCTGCTACTGACCTTACCTTTGGAG 360
QY 1558 AAGTTCTGCTGATGCTTCTTCCCTTCAATTAATGAGACTGGAAGAGCGGAGCTCA 1617
DB 361 AAGTTCTGCTGATGCTTCTTCCCTTCAATTAATGAGACTGGAAGAGCGGAGCTCA 420
QY 1618 GTCATCTCATTTGCACTGAGATGCGGAGATTTTAACTGATATTCATTTTGGAT 1677
DB 421 GTCATCTCATTTGCACTGAGATGCGGAGATTTTAACTGATATTCATTTTGGAT 480
QY 1678 AAGATTAATTTGGAACCTTTTATGGAAGAAATGAGATGTTTCCACTTATATGAC 1737
DB 481 AAGATTAATTTGGAACCTTTTATGGAAGAAATGAGATGTTTCCACTTATATGAC 540
QY 1738 CAAACAGAAATATGGAAGCAAGGATTTCTTGGAAATTTTCTAGAGTGAATCTG 1797
DB 541 CAAACAGAAATATGGAAGCAAGGATTTCTTGGAAATTTTCTAGAGTGAATCTG 600
QY 1798 CTGGCTTTTCTCATCATTTGTTTCTTATATTAATGTTCTGTTCCATTTCAAAAAAC 1857
DB 601 CTGGCTTTTCTCATCATTTGTTTCTTATATTAATGTTCTGTTCCATTTCAAAAAAC 660
QY 1858 GCTTGAAGACCAAGATGAGAAATTTTGGAAAGAGTGGCTGTTGCAATCGT 1917
DB 661 GCTTGAAGACCAAGATGAGAAATTTTGGAAAGAGTGGCTGTTGCAATCGT 720
QY 1918 TTCTTTTATAGGTTCTCTGATGAGCATGCTGAGATTCCTGATTTGTAAGTAAATC 1977
DB 721 TTCTTTTATAGGTTCTCTGATGAGCATGCTGAGATTCCTGATTTGTAAGTAAATC 780
QY 1978 CTTTCCCTCTTCCGGGTGAATATCAGACCAATGACTTCTGATAGTATTTTTC 2037
DB 781 CTTTCCCTCTTCCGGGTGAATATCAGACCAATGACTTCTGATAGTATTTTTC 840
QY 2038 CTTTCAATTAAGATGCTTGAATCCATCTCTAATCTGCAACAACCTTTTAAAG 2097
DB 841 CTTTCAATTAAGATGCTTGAATCCATCTCTAATCTGCAACAACCTTTTAAAG 900
```

```
QY 2098 GACAACTTGAACACGCTGCTGCAACAACATCAGAGAAATCAATTTCAAAATTAATAA 2157
DB 901 GACAACTTGAACACGCTGCTGCAACAACATCAGAGAAATCAATTTCAAAATTAATAA 960
QY 2158 AAAAGTTATATCAATCAATTTGTGATAGAGACTCTCTTCCCTGAACTTGGGGTT 2217
DB 961 AAAAGTTATATCAATCAATTTGTGATAGAGACTCTCTTCCCTGAACTTGGGGTT 1020
QY 2218 TTGAACAAATTAACCTTGGAGACGATATATGAACCAAGTTTCTAG 2265
DB 1021 TTGAACAAATTAACCTTGGAGACGATATATGAACCAAGTTTCTAG 1068

RESULT 12
US-09-928-175-9
; Sequence 9, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszcy, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
US-09-928-175-9

Query Match      43.9%; Score 995; DB 10; Length 1077;
Best Local Similarity 93.7%; Pred. No. 8.2e-252;
Matches 1077; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 109 CAAGTAGATGATCACTCTCTTCAATGCGCAAAAGATATTTTCCCTGAGGAATCTTACC 168
DB 1 CAAGTAGATGATGATCACTCTCTTCAATGCGCAAAAGATATTTTCCCTGAGGAATCTTACC 60
QY 169 AAGTCTTACCCCGAGCTTTTCACTGATGCGCAAGATGACTGTGGAAACGGGGCGGAC 228
DB 61 AAGTCTTACCCCGAGCTTTTCACTGATGCGCAAGATGACTGTGGAAACGGGGCGGAC 120
QY 229 GAAAGAACTGTGTGACACTAGTGTGATGCGGCGACCAATTTGGCAAGTGCATGGAAT 288
DB 121 GAAAGAACTGTGTGACACTAGTGTGATGCGGCGACCAATTTGGCAAGTGCATGGAAT 180
QY 289 GCTTAACAGGTGGCTTAAACAGAGTCTTTCTTAAACAGATCCAAATGCTGTGAC 348
DB 181 GCTTAACAGGTGGCTTAAACAGAGTCTTTCTTAAACAGATCCAAATGCTGTGAC 240
QY 349 TGCAGAGAACTGAATTTGGAATGTGAATGTGACTTAAAGTCTGTGCGATGATTTCT 408
DB 241 TGCAGAGAACTGAATTTGGAATGTGAATGTGACTTAAAGTCTGTGCGATGATTTCT 300
QY 409 AACATGTGACATTAATCTGCTCTTAAAGAAACAAATTCACAGTCTTCAAGATAAGTT 468
DB 301 AACATGTGACATTAATCTGCTCTTAAAGAAACAAATTCACAGTCTTCAAGATAAGTT 360
QY 469 TTTCATCAATTAACCAAACTTAAAGATATTTTCTCAGATTAATGCAATTAAGACATTA 528
DB 361 TTTCATCAATTAACCAAACTTAAAGATATTTTCTCAGATTAATGCAATTAAGACATTA 420
```

529 TCACGAAAGCATTTTGGATTATGATCTGCAATATTTATCTCAACCAACTGC 588
421 TCACGAAAGCATTTTGGATTATGATCTGCAATATTTATCTCAACCAACTGC 480
589 ATCAACACCCCTCAGACCTGGAAATTTCAAGACTTACAGCTTGGCTAATCTCA 648
481 ATCAACACCCCTCAGACCTGGAAATTTCAAGACTTACAGCTTGGCTAATCTCA 540
649 GATGACATCCAAATTAACAGAAATTTCAAGCTTGTTCAGGATTAATTCCTGTTT 708
541 GATGACATCCAAATTAACAGAAATTTCAAGCTTGTTCAGGATTAATTCCTGTTT 600
709 TTCTGCTATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 768
601 TTCTGCTATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
769 CCTCAACTCACTGGTGGATTTTGGAAAGCAATTAAGTATCTCAAAATTTCTACG 828
661 CCTCAACTCACTGGTGGATTTTGGAAAGCAATTAAGTATCTCAAAATTTCTACG 720
829 TTTCTGCTGCTGATTCGCTCAAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTG 888
721 TTTCTGCTGCTGATTCGCTCAAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTG 747
889 CCAAGAAACATTTTCTTCAATTAATAATTAAGAGACTGATCTGCTGCAATACG 948
748 -----CTGATCTGCTGCAATACG 768
949 ATAAAGGAGCTATCACTCACTTTTAAAGACTTGAAGCTTCAAAAGCTCAACCTG 1008
769 ATAAAGGAGCTATCACTCACTTTTAAAGACTTGAAGCTTCAAAAGCTCAACCTG 828
1009 TCATCAATCTCTTATGATCTTCAACAAGACAGTTGAAAGCTTAAACAACTTCA 1068
829 TCATCAATCTCTTATGATCTTCAACAAGACAGTTGAAAGCTTAAACAACTTCA 888
1069 TCTCTAAGCTGGAAGATGAGATTCGAATTAATAACAAGATTTCAACCACTG 1128
889 TCTCTAAGCTGGAAGATGAGATTCGAATTAATAACAAGATTTCAACCACTG 948
1129 AAGAATCTTCTCACTTATTTCAAAAAGTTGATCTGCTGCTGCTGCTGCTGCTG 1188
949 AAGAATCTTCTCACTTATTTCAAAAAGTTGATCTGCTGCTGCTGCTGCTGCTG 1008
1189 CGAATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
1009 CGAATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
1249 ATCTCTGAGA 1257
1069 ATCTCTGAGA 1077

RESULT 13
US-09-928-175-16
Sequence 16, Application US/09928175
Patent No. US20020123618A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher J.
APPLICANT: Gong, Jianhua
APPLICANT: Daugherty, Betsy
APPLICANT: Rogers, No. US20020123618A1
TITLE OF INVENTION: Leucine-rich G Protein Coupled Receptor-8 Molecules and
FILE REFERENCE: 00-1229
CURRENT APPLICATION NUMBER: US/09/928, 175
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/224,455
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 1101

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1098)
NAME/KEY: sig_peptide
LOCATION: (1)..(108)
US-09-928-175-16

Query Match 43.8%; Score 991; DB 10; Length 1101;
Best Local Similarity 93.7%; Pred. No. 9,5e-251;
Matches 1073; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

1 ATGATGTTTTCGTTTAAACATCTCTGACCTCAGATGATTAACAATGTTCTT 60
1 ATGATGTTTTCGTTTAAACATCTCTGACCTCAGATGATTAACAATGTTCTT 60
61 CTACTTCAATTCATGTTTCTGATCAATGTCAGAAATTTTGGACTGACTCAAGTACATG 120
61 CTACTTCAATTCATGTTTCTGATCAATGTCAGAAATTTTGGACTGACTCAAGTACATG 120
121 ATCACTGCTTCAGTCCAAAGATATTTTCCGTGGAAATCTTACCAAGTGTACC 180
121 ATCACTGCTTCAGTCCAAAGATATTTTCCGTGGAAATCTTACCAAGTGTACC 180
181 CGAGCTTTTCACTGATGAGCAAGATGATGTTGGAAACGGGGCGGACGAAGAACTGT 240
181 CGAGCTTTTCACTGATGAGCAAGATGATGTTGGAAACGGGGCGGACGAAGAACTGT 240
241 GGTGACACTAGTGGATGGGACCATATTTGGCAAGTGAATGCAAGGCTG 300
241 GGTGACACTAGTGGATGGGACCATATTTGGCAAGTGAATGCAAGGCTG 300
301 GCCTTAACAAGAGTCTTCTTAAACAGATATCCAAATGCTGACTGCAAGAAACT 360
301 GCCTTAACAAGAGTCTTCTTAAACAGATATCCAAATGCTGACTGCAAGAAACT 360
361 GAATGGAATGTTAAATGTTGATTAAGTCTGTCGATGATTTCTAACATGTGACA 420
361 GAATGGAATGTTAAATGTTGATTAAGTCTGTCGATGATTTCTAACATGTGACA 420
421 TTAATGCTCTTAAGAAAAAACAATCCAGTCTTCCAGATTAAGTTTCAATCAATAC 480
421 TTAATGCTCTTAAGAAAAAACAATCCAGTCTTCCAGATTAAGTTTCAATCAATAC 480
481 ACAAACTTAAAGATATTTCTCAGATTAATGATTAAGACATATCCAGAAAGCA 540
481 ACAAACTTAAAGATATTTCTCAGATTAATGATTAAGACATATCCAGAAAGCA 540
541 TTTTGGATTATGATTAATGCAATATTAATCTCAACCAACTGATCAACCTC 600
541 TTTTGGATTATGATTAATGCAATATTAATCTCAACCAACTGATCAACCTC 600
601 AGACCTGGAATATTAAGAACTTCAATCACTGATTAATGATTAATGATTAATGATTA 660
601 AGACCTGGAATATTAAGAACTTCAATCACTGATTAATGATTAATGATTAATGATTA 660
661 ATAAACGAATTTACAGCGCTTGTTCAGGATTAATGATTAATGATTAATGATTAATG 720
661 ATAAACGAATTTACAGCGCTTGTTCAGGATTAATGATTAATGATTAATGATTAATG 720
721 GTTAATACTATCTTAAGAGCTTTCCTCAAGCAATGTTGCTCAACCTCAAC 780
721 GTTAATACTATCTTAAGAGCTTTCCTCAAGCAATGTTGCTCAACCTCAAC 780
781 TGGGTGATTTGGAAGCAATTAAGTAAATGATTCACAAATTTAGCTTCTGCTG 840
781 TGGGTGATTTGGAAGCAATTAAGTAAATGATTCACAAATTTAGCTTCTGCTG 840
841 GATTCGCTACAGTGTCTTCTGCTGATTAATGATTAATGATTAATGATTAATGATTA 900
841 GATTCGCTACAGTGTCTTCTGCTGATTAATGATTAATGATTAATGATTAATGATTA 900
841 GATTCGCTACAGTGTCTTCTGCTGATTAATGATTAATGATTAATGATTAATGATTA 855

QY 901 TTTCTTCAATTAATAAATTTAGAGAACTGATCTGCTTACGAAATACGATTAACGAGCTTA 960
DB 856 -----CTGATCTGTCTAGCAATACGATTAACGAGCTTA 888
QY 961 TCACCTCACCTTTTAAAGACTTGAAGCTTCTACAAAGCTGAACCTGCTATCAATCTCT 1020
DB 889 TCACCTCACCTTTTAAAGACTTGAAGCTTCTACAAAGCTGAACCTGCTATCAATCTCT 948
QY 1021 CTATATGATCTTACAAAGAACTGTTGAAGCTTAAAGCTTAAAGCTTCAAGCTCTG 1080
DB 949 CTATATGATCTTACAAAGAACTGTTGAAGCTTAAAGCTTAAAGCTTCAAGCTCTG 1008
QY 1081 GAAAGATAGAGATTCCTCAATATTAACAACAGATGTTTCAACCTCAAGAAATCTTCTCT 1140
DB 1009 GAAAGATAGAGATTCCTCAATATTAACAACAGATGTTTCAACCTCAAGAAATCTTCTCT 1068
QY 1141 CACAT 1145
DB 1069 CACAT 1073

RESULT 14
US-09-928-175-22
Sequence 22, Application US/09928175
Patent No. US20020123618A1
GENERAL INFORMATION:
APPLICANT: Parazy, Christopher J.
APPLICANT: Gong, Jianhua
APPLICANT: Daugherty, Betsy
APPLICANT: Rogers, No. US20020123618A1
TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
FILE REFERENCE: 00-1229
CURRENT APPLICATION NUMBER: US/09/928.175
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/224,455
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 1140
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1140)
US-09-928-175-22

Query Match 36.5%; Score 825.8; DB 10; Length 1140;
Best Local Similarity 83.1%; Pred. No. 4e-207;
Matches 941; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 113 GTAGCATGATCATCTCTTCAATGCAAAAGATATTTTCCCTGGAATCTTACCAAGT 172
DB 5 GCAGTAGTGCTCTCTGCTGCCCCCAAGGATATTTTCCCTGGAATCTTACCAAT 64
QY 173 GCTTACCCCGAGCTTTTCACTGTGATGCAAGATGACTGTGGGAAACGCGGCGAGCAAG 232
DB 65 GCTTACCCCGAGCTTTTCACTGTGATGCAAGATGACTGTGGGAAACGCGGCGAGCAAG 124
QY 233 AGAAGCTGTGATGCACTAGTGATGCGGCAATATTTTGGCAACAGTCAAGAAATGCTA 292
DB 125 ACAAGCTGTGATGCACTAGTGATGCGGCAATATTTTGGCAACAGTCAAGAAATGCTA 184
QY 293 ACAAGCTGTGATGCACTAGTGATGCGGCAATATTTTGGCAACAGTCAAGAAATGCTA 352
DB 185 ATAAAGATGATGCACTAGTGATGCGGCAATATTTTGGCAACAGTCAAGAAATGCTA 244
QY 353 AAGAACTGATGATGCACTAGTGATGCGGCAATATTTTGGCAACAGTCAAGAAATGCTA 412
DB 245 GAGAAATGATGATGCACTAGTGATGCGGCAATATTTTGGCAACAGTCAAGAAATGCTA 304
QY 413 ATGAGATGATGATGCTCTTAAAGAAACAAATTCACAGCTTCCAGATTAAGTTTCA 472

DB 305 AGTAAATTAATTAATCTCTTAAAGAAACAAATTCACAGACTTCCAGTCAAGCTCTTCA 364
QY 473 TCAATTAACAAATCTTAAAGATATTTTCTTACGATTAATGCAATGACATATGCA 532
DB 365 GCAGATTAACAAATCTTAAAGATATTTTCTTACGATTAATGCAATGACATATGCA 424
QY 533 GGAAGCAATTTTGGATTAATGATTAATGCAATATTTTATGATTAATGCAACCAATGCAATCA 592
DB 425 GGAAGCAATTTTGGATTAATGATTAATGCAATATTTTATGATTAATGCAACCAATGCAATCA 484
QY 593 CAACCTCAGACCTGGAATATTTCAAGATTAATGATTAATGCAATGCAATGCAATGCAATG 652
DB 485 CCTCTCAGACCTGGAATATTTCAAGATTAATGATTAATGCAATGCAATGCAATGCAATG 544
QY 653 ACAATTCATTAACGAATTTTCAAGACCTGTTTACGGAATTAATGCAATGCAATGCAATG 712
DB 545 ACAATTCATTAACGAATTTTCAAGACCTGTTTACGGAATTAATGCAATGCAATGCAATG 604
QY 713 TGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 772
DB 605 TGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 664
QY 773 AACTCAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 832
DB 665 AACTCAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 724
QY 833 TGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 892
DB 725 TGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 784
QY 893 AAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 952
DB 785 AAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 844
QY 953 CGAGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1012
DB 845 CGAGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 904
QY 1013 CGAGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1072
DB 905 CGAGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 964
QY 1073 TAGACCTGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1132
DB 965 TAGACCTGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1024
QY 1133 ATCTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1192
DB 1025 ATCTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1084
QY 1193 TATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1245
DB 1085 TATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1137

RESULT 15
US-10-225-567A-622
Sequence 622, Application US/10225567A
Patent No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: Lifespan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burner, Glenna C.
APPLICANT: Rouher, Christine L.
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1

; SEQ ID NO 622
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-622

Query Match 32.5%; Score 735; DB 14; Length 2274;
Best Local Similarity 60.9%; Pred. No. 5,9e-183;
Matches 1263; Conservative 0; Mismatches 785; Indels 27; Gaps 3;

QY 142 GGAATTTTCCCTGGGAAATCTTACCAAGTCTTACCCCGAGCTTTCACTGTATGCG 201
DB 88 GGCATATTCCTCGTGGGAACATCACAAAGTCTTGGCTCGAGCTCTCGACTGTACGT 147
QY 202 AAGATGACTGTGGAAACGGGCGGACGAAGAACTGTGTGACACTAGTGGAGG 261
DB 148 GTGGACGACTGGGGAAATAGGCCCATGAGACMACTGTGGAGACACAAATGATGCTC 207
QY 262 ACCATATTTGGACAGTGCATGGAATGCTAAC-----GCGTG 300
DB 208 ATGCAATTTGACAAATATTTTGCAGATTACTACAAATATGACTCCCAATATCTTTGAG 267
QY 301 GCGTTAAACAGAGAGCTTTCTTAAACAGTATCCCAATGCTGTGACTGCAAGAACT 360
DB 268 GAGAAACACCTGAATGTTGGTCGGTCTGTGCGCAGTGCMAATGCTTTTGCAGAGTCTG 327
QY 361 GAATTGGAATGTAAATGTGACTTAAAGTCTGTGCGGATGATTTCTAACATGTGACA 420
DB 328 GAGCTTGACTGTGATGAAACCAATTTACAGAGCTTTCCTCAATGTGACT 387
QY 421 TTAATGCTCTTAAGAAAAAACAATTCACAGTCTTCAAGATTAAGTTTCAATCAATAC 480
DB 388 GCATGTCCTTCAAGTCAAGTAAATTAAGAAAGCTTCTCGATTGCTTCAAGATTAAT 447
QY 481 ACAAACCTAAAGATTTTCTTCAGCAATTTGATGATAGACATATCCAGAAAGA 540
DB 448 CATGATCTTCAGAGCTGTACCTGCAAAACAATTAATTAATCACTCTCACTATGCT 507
QY 541 TTTTGTGATTAATGATCTGCAATATTAATCTTCAACACAACTGATCAACCTTC 600
DB 508 TTCAAGGAGACTGAATAGCTTAACTGATTAATCTGATCAATTAAGAAATACCTTCCTG 567
QY 601 AAGCTGGAATATTCAAGACTTACATCACTGATTAATGCTGATTTTAATGACATCA 660
DB 568 AAGCCGGGTGTTTGAAGATCTTCAAGACTGATTAATGCTGATTAATGAAATATCA 627
QY 661 ATTAACAGAAATTTCAAGGCTTGTTCAGGATTAATTCCTGTTTCCGTATG 720
DB 628 CTCAGTGAATTTCCCAACACATTTTATGACTAAATTTCTTAATCTTATGCTG 687
QY 721 GTTAATTACTACTTAAGACTCTTC---CAAGCAGATGTGCCCCAAATGCTCAACTC 777
DB 688 ATGAATTAAGTCTCTCAACCGTTTACCTGAATTAACCTCTGTAACAACATGCCAATCA 747
QY 778 AACTGGGTGATTTGGAAGGCAATGAATTAAGTATCTCAAAATTTCAAGTTTCTGTG 837
DB 748 CATGTGCTGACCTTGAAGGCAACATATCAATTAATTAAGAAATTTGACTTTATTTCC 807
QY 838 TGGATTTGCTCAAGTGTGTTTCTGCTAGAAATCAATGTTGTTTGTCCAGAGAG 897
DB 808 TGCAGTAATTAATCTGTTTATGATGAGAAACAAATTAATATCACTAAATGAAT 867
QY 898 ACAATTTCTTCAATTAATAATTTAGAGAACTGATCTGTCTAGCAATAGATAGGAG 957
DB 868 ACTTTTGAACCTCTCCAGAAACTGATGATTTGATTTAGAGTAATTAAGATTAAT 927
QY 958 CTATGACCTCACTTTTAAAGCTTGAAGCTTCAAAAGCTGAACTGTATCTCAAT 1017
DB 928 CTTCACCGCTTAATTAATTAAGAGCTGAAAGAGCTGTCAATTTGATCTTCTTAAT 987
QY 1018 CCTCTATGATCTTCAAGAACAGTTGAAGTCTTAAACAATTCAGTCTAGAC 1077
DB 988 CCAATCCAGAAATTTCAAGCAACCAATTTGATTAATCTTGTCAAACTCAAGTCTCAGC 1047

QY 1078 CTGAAAGATAGATTTCCAAATTAATAACACAGAAATGTTCAACCCATGAAATCTT 1137
DB 1048 CTAGAAAGGATTGAATTTCAAAATATCCAAACAAAGAAATGTTTAACTTATGATCTC 1107
QY 1138 TCTCATATTTTTCAAAACCTTTCGATATCTGCTCTATAGCTCCCATGTCGAATAGT 1197
DB 1108 TCTCATATATTTTAAAGAAATTCAGATAGTGGATGACCAACAGTTCGAGCTGT 1167
QY 1198 ATGCCCTTGAAGGACCGGACTTCTCAATTTAGAGACCTCTTGGCTAACATATCTCAGA 1257
DB 1168 AAACCAACACTGATGAAATTTCACTTGAAGAACTCTTGGCAAGCATTTATCAGAGA 1227
QY 1258 ATATTGTCTGGTTATAGCTTTCAATTAACCTGTGGAATCTTTTGTGATTTGGCATG 1317
DB 1228 GTATTGTCTGGTTGTATCTGCAAGTTACCTGCTTTGAAACATTTTGTGATTTGACAG 1287
QY 1318 AGATCTTCAATTAAGCTGAATAATCAATCACTGATATGTCATATAAATCTTTGTTGT 1377
DB 1288 CGACCTTATATCAGGCTGAGAACAAAGCTGTATGCTCATATCTTCTCTGCTGT 1347
QY 1378 GCTGATGCTGATGAGGTTTACTGTTCTTTGTGGCAATTTGATATAAATACGA 1437
DB 1348 GCGACTCTTAATAGGAATATATTTATTCGTATGAGAGCTTTGACCTTAAGTTTGT 1407
QY 1438 GGGCAGTATCAGAAATATGCTTGTGTGATGAGACGTGACAGTGCCTCATGAGG 1497
DB 1408 GGAGAAATCAATTAAGCATGCGGAGCTGTGATGAGAGATCACTATGTCAAGTTTGA 1467
QY 1498 TTCTGGCCATGCTGTCCACGGAAGTCTGTCTGTCTACTGACTTACTTGAATTTGGAG 1557
DB 1468 TCTTGGCCATCTGTCCACAGAAATATCAGTTTATCTGTAACTTTCTGACATTTGAA 1527
QY 1558 AAGTCTGTGATGTCTTCCCTCAGTAACTTGCACCTGGAATAACGGCAGACTCA 1617
DB 1528 AATATCATCTGATGTCTATCTTTTGAATGTGAACTCTGAAATATGCAATTAAT 1587
QY 1618 GTATCTCATTTGATCTGATGAGGAGATTTTAAATGCTAATTCATTTGGAT 1677
DB 1588 ACAGTTTCAATCTCATTTGATGATCTGCTTTTATAGTGGCTTTCATTTCAATTTGAGCAT 1647
QY 1678 AAGATTAATTTGGAACCTTTTAAGGAAAAATGAAATATGTTTCCACTTAATATGAC 1737
DB 1648 AAGAAATTTTCAAAACATCATATGACCAACAAATGAAATATGCTTCTTCAATTCAGAA 1707
QY 1738 CAACAGAAATATTGAAGCAAGGATATCTTGAATTTCTTGAATTTCTTGAATTTG 1797
DB 1708 GATACAGAAATATTGAGCCCAATTAATTCAGTGCATTTTCTTGTATTAATTTG 1767
QY 1798 CTGGCTTTTCTCATATGTTGTTTCTTAATTAATTAATGTTCTGTTCAATCAAAAAC 1857
DB 1768 GCCGATTTATCATATGTTTCTTATGAAACATATGTTTATATGTTTCAATCAAGT 1827
QY 1858 GCTTGCACACCAAGAAATGAAGATTTTGAAGAGAGTGGCTTGCAAATGCT 1917
DB 1828 GCCATTAACAGCAACGAAATTAAGCAATTAAGAAATATATCTTCCAAAGCT 1887
QY 1918 TTTCTTTTATAGTGTCTGTATGCTCAATCTGATCTGATTTGTTGATTAATTAATC 1977
DB 1888 TTTTCTTTATATGTTTATTAATGATGATTAATGCTGATTAATTTTGTATGAAATTT 1947
QY 1978 CTTTCCCTCTTCCGGGTGGAATTAACAGACAAATGACTTCTGATTAATGATTTTTC 2037
DB 1948 CTTTCACTGCTTCAAGTGAATTAACAGATACCAATCACTCTTGGGTGATTTTAT 2007
QY 2038 CTTTCAAGTTAAGAGCTTTGAATCCATCTCTTAATCTTCAACAACTTTTATAG 2097
DB 2008 CTGCCATTAACAGTCTTGAACCAATCTCTAATCTGACCAACAGACCAATTTAA 2067
QY 2098 GACAAGTTGAACACCTGCTGACAAAC---TCAGAGAAATCAATTTCAAAATTA 2154
DB 2068 GAAATGATTCATGCTTTTGTATTAATTAACAGAAAGAAATCTATGAGCAAGAGT 2127

Qy 2155 AAAAAAGTTTACTACATCCATTGTTGGATAGA 2189
Db 2128 CAGAAACATATGCTCCATCATTCATCTGGTGA 2162

Search completed: November 28, 2003, 19:32:11
Job time : 732 secs

. . .

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2003, 15:41:35 ; Search time 4260 Seconds

(without alignments)
12922.450 Million cell updates/sec

Title: US-09-928-175-1

Perfect score: 2265

Sequence: 1 atgatgttttttcgtttt.....taatgaaccagtttcttag 2265

Scoring table: IDENTITY NUC

Gapop 10.0 ; Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estcov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estum:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_pro:*

26: em_gss_phg:*

27: em_gss_vit:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1249.8	55.2	2386	11	AK034592 Mus muscu
2	466.8	2838	11	AK047861 Mus muscu	
3	383.2	16.9	609	9	AV592706 AV592706
4	378	16.7	648	10	BB626659 BB626659

5	306.2	13.5	758	13	BX084029
6	296.2	13.1	841	13	BQ228832
7	290.2	12.8	648	14	CB216867
8	288.8	12.8	461	9	AV592705
9	257.4	11.4	300	9	AM307854
10	228.8	10.1	549	28	AQ394779
11	224.6	9.9	279	10	BF040869
12	190.6	8.4	857	13	BQ36143
13	190.4	8.4	779	29	BX178051
14	190.4	8.4	777	29	BX146122
15	188.6	8.3	532	28	AZ319928
16	186.2	8.2	197	12	BF942735
17	186.2	8.2	197	10	BM280297
18	184.6	8.2	235	9	AM436170
19	183.2	8.1	631	9	AL955885
20	183	8.1	408	28	AQ212353
21	181.6	8.0	811	13	BQ36385
22	179	7.9	456	10	BF021857
23	177.4	7.8	574	10	BB632910
24	177.4	7.8	1307	11	AK039086
25	174.8	7.7	1472	11	AK054129
26	171.2	7.6	744	10	BG304121
27	168.6	7.4	635	13	BQ078630
28	164.4	7.3	873	13	BQ274328
29	161.8	7.1	198	14	CA997681
30	161.6	7.1	661	13	BQ274919
31	154.4	6.8	516	10	BF158267
32	149.6	6.6	628	10	BB655906
33	145.8	6.4	307	12	BM955900
34	144.2	6.4	627	12	BM425549
35	129.6	5.7	1884	11	AK045269
36	128	5.7	720	14	CD218149
37	124	5.5	786	28	AQ745959
38	121.8	5.4	527	10	BB648375
39	118.8	5.2	593	28	AQ053279
40	116.8	5.2	544	28	AQ059136
41	106.4	4.7	696	29	BX158268
42	106	4.7	513	28	AQ059313
43	104.8	4.6	686	13	BQ284397
44	101.4	4.5	1078	29	CC244249
45	95.8	4.2	702	10	BB297224

ALIGNMENTS

RESULT 1
AK034592
LOCUS
DEFINITION
Mus musculus 12 days embryo embryonic body between diaphragm region and neck CDNA, RIKEN full-length enriched library, clone:9430012D19 product: similar to G PROTEIN COUPLED RECEPTOR AFFECTING TESTICULAR DESCENT, full insert sequence.

ACCESSION
AK034592.1 GI:26330052
VERSION
AK034592.1
KEYWORDS
HTC, CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

TITLE
JOURNAL
MEDLINE
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
JOURNAL
MEDLINE
20499374

REFERENCE AUTHORS	PUBMED 11042159
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, Y., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
TITLE	RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE PUBMED AUTHORS	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Plechmann, W., Gaasterland, T., Glass, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudl, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carinci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guernicich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Monaster, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohseuti, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE PUBMED AUTHORS	Nature 409 (6821), 685-690 (2001) 21085660 11217851
5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL AUTHORS	Nature 420, 563-573 (2002) 6 (bases 1 to 2366)
6	Adachi, J., Aizawa, K., Akimura, T., Aizawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.

FEATURES	source	Location/Qualifiers
CDS		1..2386 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM DB:9430012D19" /db_xref="taxon:10090" /clone="9430012D19" /tissue_type="embryonic body between diaphragm region and neck" /clone_11b="RIKEN full-length enriched mouse cDNA library" /dev_stage="12 days embryo" 107..1858 /note="unnamed protein product; putative similar to G PROTEIN COUPLER RECEPTOR AFFECTING TESTICULAR DESCENT (SPTR Q912Z5, evidence: PASTY, 95.4%ID, 82.9%length, match=1833)" /codon_start=1 /protein_id="BAC28765.1" /db_xref="GI:26330053" /translation="MWLLHVLITREVDPALDSSWVAPLCPRGPPCGNLTCKLPR AFRCGVDDCCNGADEDCDPTSGWTIFGVHGVNVTITLTOBGLSOYPOHCYRE NBLKCVKADLKAVERVSIINVTLSLKKKIRHLPYKVSRTLEKITYLQNRCTHIS RRAFGILNQLIVLSHNCITSLRGRFKDHLQWLILDDPIRISQSPGLNSL FFLSWGNRLBALPPTLCAQMPQLMWDLNNGIKYINSTFLTCDSTLVLELPNNQI GFVPEKTPSSKLNLGELDLSNMITKLPVHLFSDHLQKLNLSNPLVYHKNQFS KLOSLDLERLEINISTGMQPMKNSHLYLTFKRYCSYVHYRICMPSTGSISS EDLNLGIRVSVNVAIFITCVGNFLVAVSLIKAEHTTMSIKIKICACDCLMGV LPSVLPVPIKRGQYQKXALIMBSVPRFLGLFATLSTEVSULLTFLTEKPLIV FEPFSLRIGKQTAVALASIVWGLIAVPTREDYFNGFKNGVCPRLHYDQAE FGSRRYSGLIFGKFLH" 2364..2369 /note="putative" polya_signal polya_site BASE COUNT 630 a 567 c 463 g 706 t ORIGIN
Query Match	55.2%; Score 1249.8; DB 11; Length 2386; Best local Similarity 82.0%; Pred. No. 3.8e-254; Matches 1440; Conservative 0; Mismatches 317; Indels 0; Gaps 0;	
QY	32	TCAGCCTCAGATTGATTAACATGTTCTTCTACCTTCATTCATTCGTTCTGATCATGCA 91
DB	87	TCAGCGCACACCGGCGCAATGTGGCTCTCTCATCATCTCTTCTGACAGAGTCA 146
QY	92	AAGATTTCGACTGACTCAAGTAGATGATCATCTCTTCAAGCCAAAAGATATTTC 151
DB	147	AAGATTTCGACTGCTGACACAGATGATGCTCTCTGTGCCCAAGGGATATTTC 206
QY	152	CTGTGGGAATCTTACCAAGTCTTACCCGAGCTTTTCACTGTGTGCAAGATGACT 211
DB	207	CTGTGGGAATCTCACCAAAATGCTTCCCGAGCCTTTCACTGTGTGCAAGATGACT 266
QY	212	GTGGGAACGGGGCGGAGCAAGAACTGTGTGACATGTGTGATGGGAGCATATTTC 271
DB	267	GGGGGAATGTGTCCGACGAGCAAACTGTGTGACATGTGTGATGGACCATATTTC 326
QY	272	GCAAGATGATGGAATGCTTACAGCGTGCGCTTACACAGGAGTCTTCTTAAACAG 331
DB	327	GCAAGATGATGGAATGCTTACAGGAGTCTTCTTAAAGATGATGACAGAGTCTTCTTCAAGCAGT 386
QY	332	ATCCACAGACGTTTACTCTGACAGAAAGAACTGGAATGTGTAAATGTGACTTAAAGT 391
DB	387	ATCCACAGACGTTTACTCTGACAGAAAGAAAGTGTGTAAGGCTGACTTAAAG 446
QY	392	CTGTGCGCATGATTTTAAACATGTGACATTTACTGCTTTTAAAGAAACAAATTCACA 451
DB	447	CTGTGCGCAAGGTTTCAATCAACGTAACATTTACTTCTTTAAAGAAACAAATTCACA 506
QY	452	GTCCTCAGATTAAGTTTCAATCAATACACAAATTTAAAGATATTCTTCAAGCATA 511
DB	507	GACTTCAGTCAAGGCTTCTGACGATACACAGAACTCAGAAAGATATTCTTCAAGCATA 566

QY 512 ATTCATTTAGACACATATTCAGGAAGCAATTTTGGATTATGTAATCTGCAAAATATAT
 Db 567 ACTCGTCACACATCTCCAGAGAGACATTCCTGGATACATATCTACAAATATCTGT
 QY 572 ATTCACACCAACATCGATCAACACCTCAGACCCGGGAAATTTCAAAACCTTACATCAGC
 Db 627 ATTCAGCCTAACATCTCTCTCAGAGCCGGGATATTCAAACCTTGCATCAGC
 QY 632 TAACTGGCTAATCTAGATGACATCCAAATTAACAGAAATTCACAGCGCTTGTATACG
 Db 687 TTGGCTGGCTAATTTATATGACAAACCGATCCAGATCTCAGAGAAAGTCTTATAGS
 QY 692 GATTAAATTCCTGTTTCTCTCTATGTTAATTAATTAATTAATTAATTAATTAATTAAT
 Db 747 GGTAAACCTCTGTTTCT
 QY 752 AGATGTGAGCCCAATGCTCAACCTGAGTGGATTTGGAGGCAATGAAATTAAGT
 Db 807 CATGTGTCTAGATGCTCAACCTGAGTGGATTTGGAGGCAATGAAATTAAGT
 QY 812 ATTCACAAATTTCTAGTTCTGTCTGTGAGTCTGCTCAGAGTGTCTGTCTGCTAGAA
 Db 867 ACATTAACCAATCTCCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 QY 872 ATCAATTTGTTTCT
 Db 927 ATCAATTTGTTTCT
 QY 932 ATCTGTCTAGCAATTCAGTAACGAGGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 Db 987 ACCTGTCTAGCAATTCAGTAACGAGGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 QY 992 TACAAACCTGAGCTGCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 Db 1047 TCCAGAACTGAACTGCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 QY 1052 GTCTTAAACCACTTCACT
 Db 1107 GTCTTAAACCACTTCACT
 QY 1112 GAATGTTCAACCCAGTAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 Db 1167 GAATGTTCAACCCAGTAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 QY 1172 CCTATGCTCCCATGCT
 Db 1227 CCTATGCTCCCATGCT
 QY 1232 ACCTGCTGCTAACAATATCT
 Db 1287 ACCTGCTGCTAACAATATCT
 QY 1292 TTGGAATCTTTTGTGATGCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 Db 1347 TTGGAATCTTTTGTGATGCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
 QY 1352 CTATGCTCACTAAATCTTTGTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGAT
 Db 1407 CTATGCTCACTAAATCTTTGTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGAT
 QY 1412 TTGGCAATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 Db 1467 TTGGGCTCTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 QY 1472 AGAGGCTGAGTGGCGCTCATGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCT
 Db 1527 AGAGGCTGAGTGGCGCTCATGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCT
 QY 1532 TGCTACTGACCTACTGATTTGAGAGAGTCTGCTGATTTGCTGATTTGCTGATTTGCTGAT
 Db 1587 TGCTACTGACCTACTGATTTGAGAGAGTCTGCTGATTTGCTGATTTGCTGATTTGCTGAT

QY 1592 TTGACCTGGAAGAAAGCGACAGCTCAGTCAATCTCTCATTTGCTGATGCGGAGATTT
 Db 1647 TGCGCTGGGCGAGCGCGACAGCTCTGTGGCCCTGCGCAGAGCTGCGTGGGAGATTT
 QY 1652 TAATAGCTGTAATTTCCATTTTGGAAATTAAGGATTTTGGAACTTTTATGGAAGAAATG
 Db 1707 TCAATAGGCGCGCTTCCGTTTCCAGAGAGATTTTTCGGCACTTTTATGGAAGAAATG
 QY 1712 GAGTATGTTTCCCATTTATATGACCAAGAGATTTTGAAGCAAGAAAGGATTTCTC
 Db 1767 GAGTCTGCTTCCCATTTATATGACCAAGAGATTTTGAAGCAAGAAAGGATTTCTC
 QY 1772 TTGGAATTTTCTGAGT 1788
 Db 1827 TTGGGATTTTCTGAGT 1843

RESULT 2

AK047861

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

AK047861 2838 bp mRNA linear HTC 05-DEC-2002
 Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
 library, clone: C130013110 product: similar to G PROTEIN COUPLED
 RECEPTOR AFFECTING TESTICULAR DESCENT, full insert sequence.
 AK047861.1 GI:26339011
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kizunai, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Obara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nishikado, I., Pesole, G.,
 Quackenbush, J., Schriml, L. M., Strubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baderelli, R., Bahr, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Meshima, J., Mazzarelli, J., Mombere, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyono, O., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshak, B., A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S., and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 MEDLINE
 PUBMED
 21085660
 11217851

REFERENCE

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

5 (bases 1 to 2838)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furumori, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawaji, H., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submision
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

1. .2838
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM.DB:C130013110"
 /db_xref="taxon:10090"
 /clone="C130013110"
 /tissue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days embryo"
 403. .1008
 /note="unamed protein product; putative similar to G PROTEIN COUPLED RECEPTOR AFFECTING TESTICULAR DESCENT (SPTR|O91225, evidence: FASTV, 95.4%id, 82.9%length, match=1833)"

CDS

/codon_start=1
 /protein_id="BAC3177.1"
 /db_xref="GI:26319012"
 /translation="MPSDGISSSEDLANGILRVSVVIAFTCVGNFVIAVRSII KANSTTAMSIKICADCMGVLFVGVFDIKRQYOKYKALIMMSVPCRLGLFT ATLSTEVSLITLLEKPLVIFPSPNIRIGKROAVLAVIWMVGFLLAVPFR EDYFGNFGKNGCFLPHYQADDFSGRSISGLFLKRLH"
 2818. .2823
 /note="putative"
 /note="putative"
 2838
 /note="putative"
 polyA_site
 polyA_site
 729 a 688 c 561 g 860 t

BASE COUNT

ORIGIN

Query Match 20.6%; Score 466.8; DB 11; Length 2838;
 Best Local Similarity 82.7%; Pred. No. 3.2e-88;
 Matches 534; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

1143 CATTATTTCAAAAATTTCGATCTGCTCTATGCTCCCATGTCGGAATGATGCC 1202
 Db CAGTATTTGAAAACCTTTCGATCTGCTCTATGCTCCCATGTCGGAATGATGCC 407
 1203 CTTCGCGAGCGCATTTCTTCATTTGAGACCTCTGGCTAAATATCTCAGAAAT 1262
 Db GTGAGCTAGTGTATTTCTTGTCTAGAGACCTCTGGCTAAATATCTCAGAGTGC 467
 1263 TGTCTGGTTATAGCTTTCATTAATCTGCTTGGAAATCTTTTGTCTATGGATGATC 1322
 Db TGTCTGGTTATAGCTTTCATTAATCTGCTTGGAAATCTTTTGTCTATGGATGATC 527
 1323 TTTCTTTAAGCTGAAAATACATCACTGATGCTATGCTCAATCAAAATCTTTGTCGTA 1382
 Db TCTCATTAAGCTGAAAATACATCACTGATGCTATGCTCAATCAAAATCTTTGTCGTA 587
 1383 TTGCTGATGGGTGTTTCTTCTTGTGGGCAATTTTCATTAATAATACGAGGGA 1442
 Db TTGCTGATGGGTGTTTCTTCTTGTGGGCAATTTTCATTAATAATACGAGGGA 647
 1443 GTATCAGAAAGTATGCTGCTGTGTGATGAGAGCGTGCAGTCCGCTCATGAGTCTCT 1502
 Db GTATCAGAAAGTATGCTGCTGTGTGATGAGAGCGTGCAGTCCGCTCATGAGTCTCT 707
 1503 GGCATGCTGTCTCCAGGAGTCTGTCTGTCTATGACCTTCTGATTTGAGAAAGTT 1562
 Db GGCATGCTGTCTCCAGGAGTCTGTCTGTCTATGACCTTCTGATTTGAGAAAGTT 767
 1563 CTTGCTGATTTCTTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1622
 Db CTTGCTGATTTCTTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 827
 1623 CCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1682
 Db CCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 887
 1683 TTTATTTGAAATCTTTATGAGGAAATGAGATGATGATGATGATGATGATGATGATGAT 1742
 Db TTTATTTGAAATCTTTATGAGGAAATGAGATGATGATGATGATGATGATGATGATGAT 947
 1743 AGAAGATTTGGAAGCAAGGATTTCTTGGAAATTTTCTTAGGT 1788
 Db AGAAGATTTGGAAGCAAGGATTTCTTGGAAATTTTCTTAGGT 993

RESULT 3
 AV592706 609 bp mRNA linear EST 27-NOV-2001
 LOCUS AV592706 Bos taurus cartilage fetus Bos taurus cDNA clone
 DEFINITION E1CA030805, mRNA sequence.
 ACCESSION AV592706
 VERSION AV592706.1 GI:9707863
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 1 (bases 1 to 609)
 Takasuga, A., Hirotsune, S., Itoh, R., Uchizono, A., Suzuki, H., Aso, H., and Sugimoto, Y.

Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000

ESTs
 Nucleic Acids Res. 29 (22), E108 (2001)

JOURNAL MEDLINE
 PUBMED
 11713328
 Contact: Yoshihazu Sugimoto
 Animal Genetics Division

Shikrawa Institute of Animal Genetics
Odakura, Mishima, Nishi-Shikrawa, Fukushima 961-8061, Japan
Tel.: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazunugi@cocoa.com.ne.jp

Single-pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

Source

BASE COUNT	ORIGIN
128 a	137 c 145 g 189 t 10 others

Query Match	Score	DB	Length
16.98;	383.2;	9;	609;

Qy	1188	CCGAATATGTAATGACCTTCGACGAGGAGATTT--CTTCATTGAGG--ACCCCTGGGCA	1243
Db	12	CCGAATATGTAATGACCTTCGACGAGGATTTCTTTCATTGAGGACCTTTGGGCTA	71
Qy	1244	ACAAATATCTCAGAAATATTGTCT--GAGTATAGC--TTTCATTACCTGCTTTGGAAT	1295
Db	72	ACAAATATCTCAGNAGTATTGGTCTGGGGTTATAGCTTTTCATTAACCGCTTTGGAAAT	131
Qy	1300	CTTTTTCATT--GGCATGAGATCTTTTCATTAAAGCTGAAAAATCAACTCACGCTATGTC	1356
Db	132	CTTTTTCATTGGGACATGAGATCTTTTCATTAAAGCTGAAAAACAAGACTACGCTACGTC	191
Qy	1359	CATCAAAATCCCTTGTGT--GCTGATGGCTGATAGGGT--GTTACTGTGTTCTGTGGC	1412
Db	192	CATCAAAATCCCTTGTGTGGGAGCTGCTGATGGGAGGTATCTGTTCTTAAATGGC	251
Qy	1417	ATTTTC--GATATAAAAATACCGAGGGCAGTATCAGAAATATGCTTGCTGTGATGAGAG	1475
Db	252	CTTTGCGGATCTAAAGTACGCCAGGACAGTATCAGAAATGACNCGCTGTGTGATGAGAG	311
Qy	1476	CGTGAATGCGGCTCATGGGGTTCCTGGCCATGCTGTCCACCGAATCTCTGTCTGCT	1533
Db	312	CTTGAAGTGCAGGCTCTCGGGCTTCTGGCCATGCTGTCCATGAGGTCTCTGTCTCTT	371
Qy	1536	ACTGACCTACTGACTTTGGAGAAATTCCTGCTCATATGCTCCCTCAGTAACATTCG	1599
Db	372	GCTCACAATCTTGACCTCGAGAAATTCCTGCGCTGTCTTCCCTTCAGTAACATTCG	431
Qy	1596	ACCTGAAAAACGGCAGACCTCAGTATCTCATTTGCACTCGGATGCGGGAATTTTAT	1655
Db	432	TCTGGAAAAATGGCAGACCTCGCGCTCTATCTGATCTGGAATGAGGGTCTTAAAT	491
Qy	1656	AGCTGTAAATTCATTGGGAATTAAGATTTATTTGGAAACTTTATGGAAAAATGAGT	1715
Db	492	AGCTGGGACTCATATGNGAAAAAGAGATTTATTTGGAAACTTTATGAAAAAATGAGT	551
Qy	1716	ATGTTTCCACTTTATTAATGACCAACGAAGATTTGGAAGCAAGGTATCTCTT	1773
Db	552	ATGTTTCCACTTTATTAATGACCAACGAAGATTTGGAAGCAAGGTATCTCTT	609

RESULT 4	
BB626659	
LOCUS	
DEFINITION	BB626659 648 bp mRNA linear EST 26-OCT-2001
	BB626659 RIKEN full-length enriched, 12 days embryo, embryonic body
	between diaphragm region and neck Mus musculus cDNA clone
	9430012D19 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANIS

BB626659	GI:16464583
BB626659.1	
EST.	
Mus musculus (house mouse)	
Mus musculus	

TITLE
JOURNAL
COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Tsuruho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp,
url: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
wagci, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system—384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res*
10 (11), 1757-1771 (2000)
Komno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Komdo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES

Source

```

location/Qualifiers
1. 648
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9430012D19"
/tissue_type="embryonic body between diaphragm region and neck"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_1lb="Riken full-length enriched, 12 days embryo",
embryonic body between diaphragm region and neck"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15'
GAGGAGGAGGAGGATCCAGAGCGTCCTTTTCTTTTCTTTT 3', cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length b

```

cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second round cDNA was prepared with the primer adapter of sequence [5' GAGAGGAGATTCTCGATTAAATTAATTCACCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI."

BASE COUNT 191 a 164 c 137 g 156 t

Query Match 16.7%; Score 378; DB 10; Length 648;
Best Local Similarity 79.5%; Pred. No. 1.9e-69;
Matches 447; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

32 TCAGCTCAGATTGATTACATGTTCTTTCTACTTCAATTCATCGTTCTGATCATGTCA 91
87 TGGAGGCGACACCGCGCGGAGTGTGCTCTTACTTATGTCATCTTCTGACAGAGTCA 146
92 AAGATTTTGCATGACTCAAGGTAGATGATCACTCTTTCAGCCCAAAAGATATTTTC 151
147 AAGATTTTGCATGCTGACAGCAGATGATGCTCTCTGTCGCCCAAGGGTATTTTC 206
152 CCTGTGGGATCTTACCAAGTGTACCCCGACCTTTTCACTGTGATGGAGATGACT 211
207 CCGTGGGATCTTACCAAGTGTACCCCGACCTTTTCACTGTGATGGAGATGACT 266
212 GTGGGAACGGGCGGCGAGAGAACTGTGTGATCACTAGTGATGGGCGACCATATTTG 271
267 GCGGGAATGGTGGCGAGAGGACACTGTGTGATCACTAGTGATGGAGACCATATTTG 326
272 GCACAGTGCATGGAATGCTTAACAGCGTGGCTTTAACAAGAGTGTCTTTTAAACGT 331
327 GCACAGTGCATGGAATGCTTAACAGCGTGGCTTTAACAAGAGTGTCTTTTAAACGT 386
332 ATCCACAAATGCTGTGATGCAAGAACTGAATGTGAATGTGAATGTGAATGTGAAT 391
387 ATCCACAAATGCTGTGATGCAAGAACTGAATGTGAATGTGAATGTGAATGTGAAT 446
392 CTGTGCGATGATTTCTTAACATGTGATGATGATGATGATGATGATGATGATGAT 451
447 CTGTGCGAAAGTTTCATCAACGTAACTTACTATCTTTAAGAAAAAATAATCCACA 506
452 GTCTTCAATTAAGTTTTCATCAATCAACAACTTAAGAAAAAATAATCCACA 511
507 GACTTCCAGTCAAGGCTTTCAGAGATCAACAACTTAAGAAAAAATAATCCACA 566
512 ATTCATTAGACATATTCAGAGAAAGCATTTTGGATTATGTAATCTGCAATATTA 571
567 ACTGCGTACACATCTCCAGAGAGACATTCCTGGATTACATTAATCTCAATTA 626
572 ATCTCAACCAACATGATCAG 593
627 ATTTCAGCCATACTGATTAAC 648

RESULT 5 758 bp mRNA linear EST 25-APR-2003
LOCUS BX084029
DEFINITION BX084029 AGENAE Rainbow trout normalized multi-tissues library (tcac) Oncorhynchus mykiss cDNA clone tcac0004c.h.01 5prim, mRNA sequence.

ACCESSION BX084029
VERSION BX084029.1 GI:27747446
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Bukayeva; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

REFERENCE 1 (bases 1 to 758)
Govoroun,M., Guiguen,Y. and Le Gac,F.
AUTHORS Construction and primary characterization of normalized cDNA libraries in rainbow trout, Oncorhynchus mykiss
TITLE
JOURNAL
COMMENT Contact: Guiguen Y

INRA - SCRIBE
Campus de Beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us at sigenae@port@jouy.inra.fr to obtain the chromatogram of this sequence.
Plate: 0004 row: h column: 1
Seq primer: T7.
Location/Qualifiers

FEATURES
source 1..758
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcac0004c.h.01"
/tissue_type="adipose tissue, blood, brain, differentiating gonads, gills, intestinal, intestine, kidney, liver, muscle, ovary, pituitary, testis"
/dev_stage="from embryos to adults"
/lab_host="VDH10B"
/clone_lib="AGENAE Rainbow trout normalized multi-tissues library (tcac)"
/note="vector: pT7T3D-pac; Clome distribution: AGENAE Resource centre. Francois PLUMI, Francois.Plumi@jouy.inra.fr, INRA, CEA Radiobiologie et Rude du genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, France"

BASE COUNT 165 a 217 c 178 g 198 t

Query Match 13.5%; Score 306.2; DB 13; Length 758;
Best Local Similarity 65.7%; Pred. No. 3.2e-54;
Matches 446; Conservative 0; Mismatches 233; Indels 0; Gaps 0;

1491 CATGGGTTCTGTGCGATGCTGTCAACGAGTCTGTGTCTGCTACTGACTTACTTAC 1550
8 CATGGGATTCCTGGGTATGCTCTCTCAGAGATCTGTGTGCTGCTGACTGACTTACTTAC 67
1551 TTTGGAGAGTCCCGGTATGCTGTCTCCCTCAAGTAATTCGACCTGAGAAACGGA 1610
68 TGTGGAGAGTCTCGGTATGCTGTCTCTTCACTGACTGAGAGCTGAGAAACTTACA 127
1611 GACTCAGTCACTCCATTCATTCATGATGATGATGATGATGATGATGATGATGAT 1670
128 GACTCGGTGTGTGTATCTATCTGATGATGATGATGATGATGATGATGATGATGAT 187
1671 TTGAATTAAGATTAATTTTGAAGCTTTTATGAGAAAAATGAGATGATGATGATGAT 1730
188 GATGAACGAGACCTGTTGGAGACTATGATGATGATGATGATGATGATGATGATGAT 247
1731 TTATGACCAACAGAAATTAATGAGCAAGAGGATTTCTTGAATTTTCTAGGT 1790
248 CTCTGACAGACAGAGAAAGCCCACTGCTTAAGATCTTATTCGATATTTCTGGT 307
1791 GAATGCTGCTGCTTCTCATCATGATGATGATGATGATGATGATGATGATGATGAT 1850
308 GAACCTGGAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 367
1851 AAAAAACCGCTTCAGACACAGAAAGTAAGAAATGTTTGAAGAGAGTGGCTGTTC 1910
368 CAAGACTGTATCAACGCTACAGACCTGAGAGAGCTACACAGAGAGTGGCTGTTC 427
1911 AATGCTTCTTTTATATGATGTTCTGATGATGATGATGATGATGATGATGATGATGAT 1970
428 TAAAGCTTCTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487
1971 TAAATCTTCTTCTTCTTCTGCGGTGAATATCAACAGATGATGATGATGATGAT 2030
488 CAAGATACGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 547
2031 TTTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2090

Db	548	CTTATATATACCACTCAACAGTGCCTGAACCCCATCTCTACACCCCTGACTACAGCTT	607
Qy	2091	TTTAAAGACACAGTTGAAACAGCTGCTGCACAAACATCAGAGGAATTAATTTCAAAAT	2150
Db	608	CTTGAAGAAACAGGTGAGGTCTCTGCTGTGTCTGTTGGAGAGAAAGTTCAGACCTCGAAGAA	667
Qy	2151	TAATAAAAAAGTTATCT	2169
Db	668	GAAAGAAAGACCTGACCT	686
RESULT 6			
LOCUS	BQ228832	841 bp	mRNA linear EST 02-MAY-2002
DEFINITION	AGENCOURT 7522585 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6051251		
ACCESSION	BQ228832		
VERSION	BQ228832.1	GI:20410232	
KEYWORDS	EST.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 841)		
TITLE	NIH-MGC http://nigc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC/DC/DTP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: L14M13304 row: m column: 12 High quality sequence start: 87 High quality sequence stop: 603. Location/Qualifiers 1..841 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone IMAGE:6051251" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /clone_1ib="NIH MGC 72" /note="Organ: skin; Vector: pCMV-Sport6; Site 1: NCI; Site 2: Sail; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."		
FEATURES	source		
BASE COUNT	240 a 160 c 170 g 269 t	2 others	
ORIGIN			
Query Match	13.1%	Score 296.2;	DB 13; Length 841;
Best Local Similarity	67.9%	Pred. No. 4.2e-52;	
Matches 427;	Conservative 0;	Mismatches 200;	Indels 2; Gaps 1
Qy	1454	ATGCTTGTGTGATGTGAGAGACGTCGACATGCCGCTCAYGGGGTTCCTGACCATGCTT	1513
Db	1	ATGCCACAGCTGTGATGTGAGAGATCACTATGTACAGTTGTAGAGATCTTTGGCCATTCGT	60
Qy	1514	CCACGGAAGTCTGTCTGTCGCATCAGAC--TACTTGACTTTGGAGAAATTCCTGTCAT	1577
Db	61	CCACGAAATATACAGTTTACTGTTTANCAATTTTCGACATTTGAGAAAATATACATCTGCAT	120
Qy	1572	TGCTTCCCTTCAGTATACATTCGACCTGAAAACGGCAGACCTCAGTCATCTCATTTG	1631
Db	121	TGCTCATCTTTTATGATGTGTGAGACCTGAAAATGCAACAATTAAGTTCTGATTTCT	180
Qy	1632	CATTCGATGGGGGATTTTAAATAGCTGTAATTCATTTTGAATTAAGATTAATTTTGG	1693

Db	Accession	Version	LOCUS	DEFINITION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	181		CATTGGATTACGTGTTTATGAGGCTTCATTCACATTAGCAATAAGAAATTTTCAA	240							
Qy	1692		AAACTTTTATGGGAAAAATGAGTATGCTTCCACTTATATGACCAACAGAGTAT	1751							
Db	241		AAACTACTATGAGCACCATGAGTAAATGCTCCCTTCATTCAGAAATGACAGAAATAT	300							
Qy	1752		TGAGAGCAAGGGGATATCTCTTGGAATTTTCTAGGTGTAACCTGTGCTTTTCAT	1811							
Db	301		TGAGAGCCCAATTTATTCATGAGCAATTTTCTTGATTAATTTGGCCGATTTATCAT	360							
Qy	1812		CATTGCTTTTCTTAATTAATAATGTTCTGTTCCATTCAAAAACCGCTTGACAGCAC	1871							
Db	361		CATAGTTTTTCCATGAGAGCATGTTTTATGTTGTCATCAAGTCCATACAGCAAC	420							
Qy	1872		AGAAGTAAGAAATGTTTTGGAAGAGAGGGCTGTGCAAAATGTTCTTTTATAGT	1931							
Db	421		TGAAATGCGAATCAAGTTAAAAAGAGATGATCTTGCCAAAGTTTTTCTTTATAGT	480							
Qy	1932		GTTCTGATGCCATCTGCTGAGATTCCTGTAATTTGATGAATAAATCCTTTCCTCCG	1991							
Db	481		ATTACTGATGCAATTAATGCTGATACCCATTTTTTGATGAAATTTCTTTCACGCTTCA	540							
Qy	1992		GGTGAATACCGACACCAATGACTCTGAGTATGATTTTTTCTTCAGTTAACAG	2051							
Db	541		GGTGAATACCGAGTACCAATGACTCTGAGTATGATTTTTTCTTCAGTTAACAG	600							
Qy	2052		TGCTTTGAATCCATCCTCTAATCTTCA	2080							
Db	601		TGCTTTGAATCCATCCTCTAATCTTCA	629							
RESULT 7											
CB216867			648 bp	mRNA	linear	EST 06-FEB-2003					
LOCUS			NISC.ng09g08.y1.NICH.D_HS.Ut2	Homo sapiens	cdna	IMAGE:5938742					
DEFINITION			5', mRNA sequence.								
ACCESSION			CB216867								
VERSION			CB216867.1	GI:28265055							
KEYWORDS			EST.								
SOURCE			Homo sapiens (human)								
ORGANISM			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
REFERENCE			1 (bases 1 to 648)								
AUTHORS			NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .								
TITLE			National Cancer Institute, Cancer Genome Anatomy Project (CGAP),								
			Tumor Gene Index								
			Unpublished								
JOURNAL			Contact: Robert Strausberg, Ph.D.								
COMMENT			Email: cgapbs-remail.nih.gov								
			CDNA Library Preparation:								
			CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL								
			DNA Sequencing by: National Institutes of Health Intramural								
			Sequencing Center (NISC)								
			Clone distribution: NCI-CGAP clone distribution information can be								
			found through the I.M.A.G.E. Consortium/LLNL at:								
			info@image.llnl.gov								
			Plate: LLAM13167	row: M	column: 15						
			Seq primer: M13RPL reverse primer (ABI).								
FEATURES			Location/Qualifiers								
source			1..648								
			/organism="Homo sapiens"								
			/mol_type								

from normal endometrial tissue (mid-secretory phase, cycle day 23). Average insert size 1.6 kb. Library constructed by RegGen (Invitrogen Corporation)."

BASE COUNT

180 a 112 c 133 g 223 t

ORIGIN

Query Match 12.8%; Score 290.2; DB 14; Length 648;
Best Local Similarity 65.5%; Pred. No. 7.7e-51;
Matches 424; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

QY 1255 AGAATATTTGCTGGGTATAGCTTTCATTAACCTGCTTGGAAATCTTTTGTGATGGC 1314
DB 2 AGAGTATTTGCTGGGTATAGCTTTCATTAACCTGCTTGGAAATCTTTTGTGATGGC 61
QY 1315 ATGAGATCTTTCATTAAGCTGAAATACAACTCAGCGTATGCTCAAAATCTTTG 1374
DB 62 ATGCACTTATATACAGCTTTCAGAAACAGCTTATGCAATCAATTTCTCTGC 121
QY 1375 TGTGCTGATTCCTGATGAGGTGTTTACTTGTCTTTGTTGGCATTTTCATATAATAC 1434
DB 122 TGTGCGAGCTGCTTATGGAATATATTTATGATGAGGCTTTGACCTAAAGTTT 181
QY 1435 CGAGGCAATGATGAAAGTATGCTTGTGATGAGAGCGTGAAGCTGCGCTCATG 1494
DB 182 CGTGAGAAATACAAATAGCATGCGCAGCTGTGAGAGTACTCATTTGCTGTA 241
QY 1495 GGGTCTGCGCAATGCTGTCACCGAAGTCTGTTCTGCTACTGACTTACTTGG 1554
DB 242 GATCTTTGGCCATCTGTCACAGAAATACAGTTTATCTTGAACATTTCTGACATTTG 301
QY 1555 GAGAAATTCCTGATCTGCTTCCCTTCAGTATGATGAGCTGAGAAACGCGAGACC 1614
DB 302 GAAATATACATGCAATGCTTCTTAAAGTGTGAGACTGAGAAATGACAGAA 361
QY 1615 TCGATATCTCATTTGATCTGATGAGCGGATTTTATATGCTGATTTCAATTTTG 1674
DB 362 ATTACAGTTCTGATCTCATTTGATGATTAAGTGTGAGCTTTCATTTCAATGAG 421
QY 1675 AATAAGATATTTTGGAACTTTTATGGAATAATGAGATATGTTCCACTTTATAT 1734
DB 422 AATAAGATATTTTCAAACCTATGATGAGCAACAGATATGCTTCTTCAATCA 481
QY 1735 GACCAACAGAGATATTTGAGCAAGGATTTCTTGGAAATTTTCTTGAAGTGAAC 1794
DB 482 GAAGATACAGAAATATTTGAGCCAGATTTATTCAGTGGCAATTTTCTTGATTAAT 541
QY 1795 TTGCTGCTTTCTCATCTATGTTTCTTATATTAATGTTCTGTTGATTTCAAAA 1854
DB 542 TTGCGCGCATTTATCATCATAGTTTCTTATGGAAGCATGTTTATATGTTTCA 601
QY 1855 ACCGCTTTCAGACACAGAAATGATTTGTTTGAAGAGGT 1901
DB 602 AGTGCCATACACCACTGAAATACGATCAAGTTAAAAAGAGAT 648

RESULT 8

AV592705/c

LOCUS AV592705 461 bp mRNA linear EST 27-NOV-2001
DEFINITION AV592705 Bos taurus cartilage fetus Bos taurus cDNA clone
E1CA003B08 3', mRNA sequence.

ACCESSION AV592705
VERSION AV592705.1 GI:9707862
KEYWORDS EST.

SOURCE

Bos taurus (cow)
Bos taurus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 461)
Takaue, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000

JOURNAL Bovine ESTs
MEDLINE Nucleic Acids Res. 29 (22), E108 (2001)
PUBMED 21570554
11713328

COMMENT

Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan.
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cococ.ocn.ne.jp

Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

1..461
Location/Qualifiers

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1CA003B08"
/issue_type="cartilage"
/dev_stage="fetus"
/lab_host="DH10B"

/clone_1ib="Bos taurus cartilage fetus"
/note="vector: pZ1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

BASE COUNT 134 a 104 c 109 g 109 t 5 others

ORIGIN

Query Match 12.8%; Score 288.8; DB 9; Length 461;
Best Local Similarity 87.0%; Pred. No. 1.5e-50;
Matches 314; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1433 ACCGAGGCGATGATGAAAGTATGCTGCTGATGAGAGCGTGCAGTCCGCTCA 1492
DB 460 ACCGAGGCGATGATGAAAGTATGCTGCTGATGAGAGCGTGCAGTCCGCTCA 401
QY 1493 TGGGGTCTGCGCATGCTGTCACCGAAGTCTGTTCTGCTACTGACTTACTT 1552
DB 400 TGGGCTCTGCGCATGCTGTCACCGAAGTCTGTTCTGCTACTGACTTACTT 341
QY 1553 TGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1612
DB 340 TGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 281
QY 1613 CCTCAGTCAATCTCATTTTGCATCTGATGAGCGGATTTTATATGCTGATTTT 1672
DB 280 CCTGCGCTCTCATCTGATCTGATGAGCGGATTTTATATGCTGATTTT 221
QY 1673 GGAATAGGATTTTATGGAATCTTTATGGAATAATGAGATGTTTCCACTTAT 1732
DB 220 GGAATAGGATTTTATGGAATCTTTATGGAATAATGAGATGTTTCCACTTAT 161
QY 1733 ATGACCAACAGAGATATTTGGAAGCAAGGATTTCTTGGAAATTTTCCATG 1792
DB 160 ATGACCAACAGAGATATTTGGAAGCAAGGATTTCTTGGAAATTTTCCATG 101

RESULT 9

AM307854

LOCUS AM307854 300 bp mRNA linear EST 25-APR-2001
DEFINITION AM307854 4BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AM307854
VERSION AM307854.1 GI:6720217
KEYWORDS EST.

SOURCE

Bos taurus (cow)
Bos taurus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 300)
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
 G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
 Pettea,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and
 Keefe,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 JOURNAL libraries and construction of a gene index for cattle
 MEDLINE Genome Res. 11 (4), 626-630 (2001)
 PUBMED 21180013
 COMMENT 11282978
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCATGACCAT
 BACKWARD: GTTTCACAGCAGACG
 Plate: 142 row: H column: 17
 Seq primer: ATTAGTGACACTATAG.
 Location/Qualifiers
 1..300
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_1ib="MARC 4BOV"
 /note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from day 20 and day 40
 embryos."

BASE COUNT 62 a 69 c 74 g 95 t
 ORIGIN

Query Match 11.4%; Score 257.4; DB 9; Length 300;
 Best Local Similarity 91.3%; Pred. No. 6.6e-44;
 Matches 273; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1203 CTTCGACGACGCAATTCCTTCATTGAGGACCTCTTGCTTAACAATATCTCAGATATT 1262
 DB 2 CTTCGACGACGCAATTCCTTCATTGAGGACCTCTTGCTTAACAATATCTCAGATATT 61
 QY 1263 TGTCTGGTTATAGCTTTTCAATACCTGCTTTGGAATCTTTTGTGATGGCATGAGATC 1322
 DB 62 TGGCTGGTTATAGCTTTTCAATACCTGCTTTGGAATCTTTTGTGATGGCATGAGATC 121
 QY 1323 TTTCATTAAGCTGAATAATACAACCTACGCTATGTCATCAAAATCCTTTGTTGCTGA 1382
 DB 122 TTTCATTAAGCTGAATAATACAACCTACGCTATGTCATCAAAATCCTTTGCTGCGGA 181
 QY 1383 TTGCTGATGGGTCTTACTGTTCTTGTGCTTTGCTGATTTTGAATATAATACCGAGGCA 1442
 DB 182 CTGCTGATGGGTCTTACTGTTCTTGTGCTTTGCTGATTTTGAATATAATACCGAGGCA 241
 QY 1443 GTATACGAGATATCCTTCTGCTGATGAGAGCGGCGATGCCGCTCATGGGCTTC 1501
 DB 242 GTATACGAGATATCCTTCTGCTGATGAGAGCGGCGATGCCGCTCATGGGCTTC 300

RESULT 10
 A0394779 549 bp DNA linear GSS 06-MAR-1999
 LOCUS CITR1-El-2548P17.TR CITR1-El Homo sapiens genomic clone 2548P17,
 DEFINITION genomic survey sequence.
 ACCESSION A0394779
 VERSION A0394779.1 GI:4365802
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 549)
 AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., Shizuya,H., Simon,M. and
 Venter,J.C.
 TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
 JOURNAL Map Building
 MEDLINE Unpublished
 PUBMED Other GSSs: CITR1-El-2548P17.TF
 Contact: Shaying Zhao, William Niernan, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbo@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..549
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="2548P17"
 /sex="male"
 /cell_type="sperm"
 /clone_1ib="CITR1-El"
 /note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
 Caltech Human BAC Library D"

BASE COUNT 152 a 110 c 96 g 189 t 2 others
 ORIGIN

Query Match 10.1%; Score 228.8; DB 28; Length 549;
 Best Local Similarity 98.3%; Pred. No. 7.9e-38;
 Matches 230; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1144 ATTTATTCAAAATCTTGATGATGCTGCTGATGCTCCCATGTCGATATGATGCCC 1203
 DB 226 ACTATATTCAAAATCTTGATGATGCTGCTGATGCTCCCATGTCGATATGATGCCC 285
 QY 1204 TTGACGACGCAATTCCTTCATTGAGGACCTCTTGCTTAACAATATCTCAGATATT 1263
 DB 286 TTGACGACGCAATTCCTTCATTGAGGACCTCTTGCTTAACAATATCTCAGATATT 345
 QY 1264 GTCTGGTTATAGCTTTTCAATACCTGCTTTGGAATCTTTTGTGATGGCATGAGATC 1323
 DB 346 GTCTGGTTATAGCTTTTCAATACCTGCTTTGGAATCTTTTGTGATGGCATGAGATC 405
 QY 1324 TTTCATTAAGCTGAATAATACAACCTACGCTATGTCATCAAAATCCTTTGTTGT 1377
 DB 406 TTTCATTAAGCTGAATAATACAACCTACGCTATGTCATCAAAATCCTTTGTTGT 459

RESULT 11
 BF040869 279 bp mRNA linear EST 10-OCT-2000
 LOCUS BP25001A10B3 Soares normalized bovine placenta Bos taurus cDNA
 DEFINITION Clone BP25001A10B3 5', mRNA sequence.
 ACCESSION BF040869
 VERSION BF040869.1 GI:10757924
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 279)
 AUTHORS Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson,
 J.H.

TITLE Bovine ESTs
JOURNAL Unpublished
COMMENT Contact: Lewin, H. A.
W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Funding for cattle EST sequencing was provided by the USDA National
Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
from Washington University Genome Center. Vector Trim1 g:
Cross_match from Washington University Genome Center PHRAP suite.
This sequence is vector free and at least 200 bp in length.
PCR primers
FORWARD: TAATACGACTCACTAAGG
BACKWARD: ATTAACCTCACTAAG
Insert Length: 279 Std Error: 0.00
Plate: BP25001A10 row: B column: 3
Seq primer: AGCGGATACAACTTCACACGGA
High quality sequence stop: 279.
Location/Qualifiers
1. 279
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP25001A10B3"
/sex="female"
/lab_host="DH10B"
/note="Organs: Soares normalized bovine placenta"
/note="Organ: placenta; Vector: pT73Pac; Site 1: EcoRI;
Site 2: NotI; The cDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806."

BASE COUNT 51 a 75 c 70 g 83 t

Query Match 9.9%; Score 224.6; DB 10; Length 279;
Best Local Similarity 87.8%; Pred. No. 5.9e-37;
Matches 245; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1306 GTGATGGAGATGATCTTTTATTAAAGCGAAATCACTCAAGCTATGTCATGAA 1365
1 GTGATGGAGATGATCTTTTATTAAAGCGAAATCACTCAAGCTATGTCATGAA 60
Db 1 ATCCTTTGTTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1425
QY 1366 ATCCTTTGTTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1425
Db 61 ATCCTTTGTTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 120
QY 1426 ATAAATACCGAGGCGATATCAGAAAGTATGCTGCTGAGAGAGAGCTGCAATGC 1485
121 CTAAAGTACCGAGGCGATATCAGAAAGTATGCTGCTGAGAGAGAGCTGCAATGC 180
QY 1486 CGCTCATGAGGCTTCTGAGCATCTGTCACAGAAAGTCTGTTGCTACTAGCACTAC 1545
Db 181 CGCTCATGAGGCTTCTGAGCATCTGTCACAGAAAGTCTGTTGCTACTAGCACTAC 240
QY 1546 TTGACTTTGAGAGAGTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1584
Db 241 TTGACTTTGAGAGAGTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 279

RESULT 12
BUS66143 857 bp mRNA linear EST 28-NOV-2002
LOCUS BUS66143 60358554.f1 CSECHN72 Gallus gallus cDNA clone CSE5141a11 5', mRNA
DEFINITION BUS66143
ACCESSION BUS66143.1 GI:25874144
VERSION
KEYWORDS EST.

SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 857)
AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Boesch, E.,
Fong, M.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
JOURNAL MEDLINE
PUBMED 12445392
CONTACT: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 857
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton line 151"
/db_xref="taxon:9031"
/clone="CSE5141a11"
/sex="female"
/tissue_type="cerebrum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSECHN72"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones; cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research
(1996): 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT 259 a 191 c 167 g 240 t

Query Match 8.4%; Score 190.6; DB 13; Length 857;
Best Local Similarity 54.2%; Pred. No. 1e-29;
Matches 445; Conservative 0; Mismatches 354; Indels 22; Gaps 2;

QY 108 TCAAGGTAGATATATCATCTTCATGCGCAAAAGATATTTCCCTGCGGAATCTTAC 167
23 TCCTGTTTCCAGAGGGGGTCTGCTGCGCCCTTGGCTACTTCTTCCCTGCGCAATATCAC 82
Db 168 CAAGTGTATACCCCGAGCTTTTCACTGTGATGCGCAAGATGATGCTGCGGACCGAGCGGA 227
83 AAGTGTATACCCCGAGCTTTTCACTGTGATGCGCAAGATGATGCTGCGGACCGAGCGGA 142
QY 228 CGAAGAGACTGTGTGATCACTAGTGAATGGGACCAATATTTGGCAAGTGAAGAA 287
Db 143 TGAAGCAACTGTGATGATCACTAGTGAATGGGACCAATATTTGGCAAGTGAAGAA 202
QY 288 TGCTAACA-----GCGTGGCTTAACACAGAGAGCTTTCTAATAACA 329
Db 203 AGACAAATATAAATCCCTGTAGCATTTTCAAGCAAGACACCAAGATGCTTGGCGGTGC 262
QY 330 GTATCCCAATGCTGTGATGCAAGAAACTGAATTGATGTGTAATGTGTAACTTTAA 369
Db 263 TGATCCAGTGAATGTAAATGCAAGGGCTGAGAGGCTTCTGTGATGCTGCCAAACTGCG 322

QY 390 GCTGTGCCATGATTTCTTAAAGATGACATTAAGTCTCTTAAAGAAAAAATAATCCA 449
DB 323 TGATGTCCCTTTGGTCCCTTCAAAATACACCATGCTCTTCAAGAAAGCTTCTAAG 382
QY 450 CAGCTTCCAGATTAAGTTTCAATCAATACAAATTTAAAGATTTCTTCAAGCA 509
DB 383 GAAACTTATGCTACGTTTCAAGAAATACCAAGACTTAAATCTGATCTTCAAGCA 442
QY 510 TAAATGATTAAGACATATCCAGAAAGCAATTTTGGATTAATGATCTGCAATATTT 569
DB 443 TAAATTAATGAGACTGTGTCCAAACAGCTTTCAAGGTTTATACAACTTGACAAATTT 502
QY 570 ATATCTCAACCAATGTCATACACCTTCAAGCTGGAATTTCAAGATCTTACATCA 629
DB 503 ATACCTAGATTAACCAAGATTAACCACTTGAAGCTGTCTTGAAGATCTTCAATTA 562
QY 630 GCTAATCTGCTAATTCAGATGACATTCATTAACCAATTTTCAAGGCTTGTATAC 689
DB 563 ACTGGAATGCTGATTAATTAAGAAATTAAGATTAATGATCTCTCATCAACATTTTA 622
QY 690 GGAATTAATCTGTTTCTGTCTGTATGATTAATTAATTAATTAATTAATTAATTAAT 746
DB 623 TGGACTCAATCACTATATCTCTTGAATATGATTAATTTCTTGTCTATTTGCTG 682
QY 747 -CAAGCAGATGTGCCCCAAATGCTCACTCACTGCTGATTTGGAAGCAATGAA 805
DB 683 ACAACCTCTGTGCAATATATGCAAGATGAACTGAGCTGAGCTTGAAGCAACCA 742
QY 806 TAAATGATCTCAAAATTTCTACGTTTCTGTGTGCGATTCCTCAAGTGTCTTTCTGC 865
DB 743 TTCTACTTAAAGAAATGCTCACTTCACTCTGACAGCACTTAACTGTTTGGGTGATGA 802
QY 866 CTAGAAATCAATGTTGTTTCTGTCAGAGAAAGACATTTCT 906
DB 803 GACAAATTAATCGGCTTAAATGAGAAACAGTTTCTCT 843

RESULT 13
BX178051 739 bp DNA linear GSS 13-MAR-2003
LOCUS BX178051
DEFINITION Danio rerio genomic clone DKEY-176C8, genomic survey sequence.
ACCESSION BX178051
VERSION BX178051.1 GI:28009822
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 739)
Humphrey, S.J., Huckle, E. and Durham, J.L.
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphrey@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 176C8, 176C8 is
part of the Daniokey BAC library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/
location/Qualifiers
1. 739

FEATURES
Source
1. 739
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-176C8"
/issue_type="Testis"
/note="Vector pIndigoBAC-536"
BASE COUNT 188 a 141 c 133 g 277 t
ORIGIN

Query Match 8.4%; Score 190.4; DB 29; Length 739;
Best Local Similarity 66.1%; Pred. No. 1.1e-29;
Matches 275; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 1377 TGGTATGCTGATGAGGTTTACTTGTCTTGTGTCATTTTCAATTAATACCG 1436
DB 220 TGGCGATGCTTATGGAATCTTACCTTTCTTCTGGGGAATTTTATGTAATTCGG 279
QY 1437 AGGCGATATCAAGATATCTTCTGTGTGATGAGAGGCTGACGCTCCATGGG 1496
DB 280 TGGCGATATCAACCGAACCTCTGATCTGATGAGACAGGCTGTAACCAATTTGG 339
QY 1497 GTTCTGCGCATGCTGTCACGGAAGCTGTGTTCTGTCTGACCTGACTTGA 1556
DB 340 CTTCCTGCGCATGCTGTCACGGAAGCTGTGTTCTGTCTGACCTGACTTGA 399
QY 1557 GAAATTCCTGCTATGCTTCTTCCCTTCACTTAATTTGACCTGGAAGAGGAGCTTC 1616
DB 400 GAAATTCCTGCTATGTTTCTTCCCTTACCAATTTGAGACCTGCAAGCTGAGAGGT 459
QY 1617 AGTCATCTCATTTGCTGATGAGTGGCGGATTTTAAATGCTGTAATTCATTTGGA 1676
DB 460 GCTGATTTCTGCTTCACTGTTTCTGAGATTTGTCATAGCAGCTGTCCCTTTAAA 519
QY 1677 TAAATATTTTGAATCTTATGGAATTAAGATTAATGATTAATTAATTAATTAAT 1736
DB 520 CGAAGCTGTGTTGGGATTAATTAACGCGAGAAATGAGTTGTTCTCTCCATCCGA 579
QY 1737 CCAACAGAGATATTTGGAAGCAAGGATTTCTCTGGAATTTCTAGGTGGA 1792
DB 580 TAGCGTGAAGAACCAAGCGCAAGATATCTCACTGGAATTTCTTGTGTTGA 635

RESULT 14
BX146122 777 bp DNA linear GSS 28-JAN-2003
LOCUS BX146122
DEFINITION Danio rerio genomic clone DKEY-108122, genomic survey sequence.
ACCESSION BX146122
VERSION BX146122.1 GI:27977480
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 777)
Humphrey, S.J., Huckle, E. and Durham, J.L.
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphrey@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 108122, 108122
is part of the Daniokey BAC library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/
location/Qualifiers
1. 777

FEATURES
Source
1. 777
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-108122"
/issue_type="Testis"
/note="Vector pIndigoBAC-536"
BASE COUNT 200 a 147 c 140 g 290 t
ORIGIN

Query Match 8.4%; Score 190.4; DB 29; Length 777;
Best Local Similarity 66.1%; Pred. No. 1.1e-29;
Matches 275; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 1377 TGGTATGCTGATGAGGTTTACTTGTCTTGTGTCATTTTCAATTAATACCG 1436
DB 235 TGGCGATGCTTATGGAATCTTACCTTTCTTCTGGGGAATTTTATGTAATTCGG 294
QY 1437 AGGCGATATCAAGATATCTTCTGTGTGATGAGAGGCTGACGCTCCATGGG 1496

```

Db      295 TGGCGATACACCGGAGCGCTGTGATCTGATGAGACAGGCTTGATGCCGAACATTTGG 354
QY      1497 GTTCTGGCCAGCTCTGTCCACGAGAGTCTCTGTTCTGCTACTGACTTATGACTTTGGA 1556
DB      335 CTTTCCTGGCAGCTCTCTCCCTGAGAGTCTGAGTCTCTCTCTCCCTCACTTGAACATTA 414
QY      1557 GAAGTCTGGGATTTGTTCTCCCTTCAGTACATTCGACCTGGAAAACGAGACCTTC 1616
DB      415 GAATTCCTGTGATTTGTTTCCCGTTTACCTTTGAGACCTGCCAAGCTGACAGCGGT 474
QY      1617 AGTCACTCTCATTTGATCTGATGAGGAGATTTTAAATAGTCTATTCATTTGGA 1676
DB      475 GCGATTTCTGCTTATCTGATCTGTTCTCGGATTTGATGACAGCTGTTCCCTGTTAA 534
QY      1677 TAAGATTTATTTGGAACTTTTATGGAATAATGAGATGTTTCCCTTATTTATGA 1736
DB      535 CGAAGCTGTGTTGGGAAATTTATACGCGAGAAATGAGTGTGTTCCCTCTCCACCTCGA 594
QY      1737 CCAACAGAAATATTTGAAAGGAGATTTCTCTGTAATTTCTCTAGTGTGA 1792
DB      595 TAGGCTGGAGAAACCAACGCGCAAGGATCTCCTGAGATTTCTTGTGTTGA 650

```

```

RESULT 15
A2319928/c 532 bp DNA linear GSS 29-SEP-2000
LOCUS      1M0039D13R Mouse 10kb plasmid UUC1M library Mus musculus genomic
DEFINITION clone UUC1M0039D13 R, genomic survey sequence.
ACCESSION  A2319928
VERSION     A2319928.1 GI:10371193
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 532)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
            and Wright,D., Weisb,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished
            Contact: Robert B. Weisb
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert length: 10000 Std Error: 0.00
            Plate: 0039 row: D column: 13
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 532.
            Location/Qualifiers
                1..532
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="UUC1M0039D13"
                    /sex="Male"
                    /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F--"
                    /clone_lib="Mouse 10kb plasmid UUC1M library"
                    /note="Vector: PWD42nv; Purified genomic DNA from M.
                    musculus C57BL/6J (male) was obtained from the Jackson
                    Laboratory Mouse DNA Resource
                    (http://www.jax.org/resources/documents/dnares/). The DNA
                    was hydrodynamically sheared by repeated passage through a
                    0.005 inch orifice at constant velocity. The sheared DNA
                    was blunt end-repaired with T4 DNA polymerase and T4

```

polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 184 a 120 c 95 g 133 t

ORIGIN

Query Match 8.3%; Score 188.6; DB 28; Length 532;
 Best Local Similarity 87.7%; Pred. No. 2.6e-29;
 Matches 206; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

```

QY      1143 CATTTATTTCAAAACTTTGATGCTGCTCTATGCTCCCATGTCGGAATATGATGCC 1202
DB      511 CAGTTATTTGAAAACCTTTCGATCTGCTCTATGCTCCCATGTCGGAATCTGATGCC 452
QY      1203 CTGACGAGCGCATTTCTTCAATTGAGACCTCTTGCTAACAAATATCTCAGATATT 1262
DB      451 GTGACGTATGATGATTTCTGCTGAGAGACCTCTTGCTAACGATATCTCAGAGTGC 392
QY      1263 TGTCTGGTTATAGCTTTTATTAAGCTCTTGGAATCTTTTGTTCATTTGATGATGC 1322
DB      391 TGTCTGGTTATAGCTTTTATTAAGCTCTTGGAATCTTTTGTTCATTTGATGATGC 332
QY      1323 TTTTCATTAAGCTGAAATATACACTGAGCTATGTCATCAAAATCCTTGTGTG 1377
DB      331 TCTCATTAAGCTGAGAAATACACTGAGCTATGTCATCAAAATCCTTGTGTG 277

```

Search completed: November 28, 2003, 19:17:04
 Job time : 4266 secs

Db 301 CATTAGCTGATACCATTTTGTAGTAAATTTCTTACCTGCTTCAGTGAATAC 360
Qy 2003 CAGACACATGACTCTCTGATAGTATTTTCTTCCAGTTACAGCTTTGATC 2062
Db 361 CAGGTACCATTAACCTCTGGGTAGTATTTTATCTGCCCCATTAACAGCTTTGAAC 420
Qy 2063 CATTCCCTTACTCTTCACACCACTTTTAAAGCAAGTTGAACACCTGCTGACA 2122
Db 421 CAATCTCTTACTCTGACCAAGACCATTTAAAGAAATGATTCATGCGTTTGTATA 480
Qy 2123 AAC--TCAGAGGAATCAATTTTCAAAATTAATAAAAGTTTATCTACATCCATG 2179
Db 481 ACTACACACAAAGAAATCTATGACAGCAAGCTCAGAAACATATGCTCATCATCA 540
Qy 2180 TGTGATAGA 2189
Db 541 TCTGGGTGGA 550

RESULT 2
US-08-487-886-1
; Sequence 1, Application US/08487886
; Patent No. 5744448
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweickhardt, Rene Lynn
; APPLICANT: Cheng, Shirley Vui Yen
; APPLICANT: Nugent, No. 5744488en Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan P. Williams,
; ADDRESSEE: Ares-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Masell via Kermit to IBM MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: 15-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
; TELEFAX: (617) 723-8923
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2179
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Testis
; IMMEDIATE SOURCE:
; LIBRARY: Igtil cDNA library, Clontech #HL1010b
; CLONE: pHPShr11-11, pHPShr15-6
; FEATURE:

NAME/KEY: protein coding region
; LOCATION: 75 to 2159
US-08-487-886-1

Query Match 2.8%; Score 63.6; DB 1; Length 2179;
Best local Similarity 47.3%; Pred. No. 3.1e-08;
Matches 192; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

Qy 1166 ACTGCTCTATGCTCCCATGTCCTCAATATGATAGCTCTTACAGAGGAGCATTTCTCAT 1225
Db 1081 ACTTATGACATAGAGTGTGACGTAAGCTGCTCCCTTAAGCCAGATTCACACCAT 1140
Qy 1226 TTGAGGACCTCTGCTTAACAATATCTCAGATATTTGCTGGTTATAGCTTTATTA 1285
Db 1141 GTGAGATATCATGGGTGTCACACATCTCAGAGCTCTGATATGTTTATCAGATCTCG 1200
Qy 1286 CCTGCTTGGAAATCTTTTGTCTATGTCAGATGAGATCTTTCATTAAGCTGAATAACAA 1345
Db 1201 CCATCACTGGAAACATATAGTGTATGATCTTAACTACAGCCAAATTAACCTACAG 1260
Qy 1346 CTCAGCCTATGTCATCAAAATCTTTGTTGCTGATGCTCCATGAGGGTGTACTGT 1405
Db 1261 TCCCCAGGTTCTTATGTCACACCTGCTTGTCTGATCTGCAATGAAATCTTACCTGC 1320
Qy 1406 TCTTGTGACATTTTCGATATATAATACAGAGGCGAGTATGAGATGCTTGTCTGT 1465
Db 1321 TGCTATTTGATCATAGTATATATTCATACCAAGAGCCAAATATCAATATGATGACT 1380
Qy 1466 GGATGAGAGCGTGCAGTGCCTGCTCATGAGGTTCTTGGCATGCTGTCCAGGAAGTCT 1525
Db 1381 GCGAAACTGGGGGAGGCTGTGATGTCGCTGCTTTTCACTGTGCGCAGTGAAGTCT 1440
Qy 1526 CTGTTGCTACTGACCTACTGCTTGTGAGAGAGTTCTGAGCT 1571
Db 1441 CAGTCTACACTGTGACAGCTATACCTTGGAAAGATGCAATACAT 1486

RESULT 3
US-08-531-070A-1
; Sequence 1, Application US/08531070A
; Patent No. 5851768
; GENERAL INFORMATION:
; APPLICANT: de la Chapelle, Albert
; APPLICANT: Altomaki, Kristina
; APPLICANT: Hublaniemi, Ilpo
; TITLE OF INVENTION: Method For Diagnosis Of Ovarian Dysgenesis
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,070A
; FILING DATE: 20-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/32879
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

LENGTH: 2179 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-531-070A-1

Query Match 2.8%; Score 63.6; DB 2; Length 2179;
Best Local Similarity 47.3%; Pred. No. 3.1e-08;
Matches 192; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

1166 ACTGCTCTATGCTCCCATGTCGAATATGATGACCTTGAACGAGGCAATTTCTTAT 1225
1081 ACTATGCAATGAAGTGTATACGACCTGCTCCCAAGCCAGATGCAATCAACCAT 1140
1226 TTGAGGACCTTGGCTTACAAATATCTCAGAAATTTGTGGGTTTATGCTTATTA 1285
1141 GTGAAGATATCATGGGGTACAAACATCTCAGAGTCCATATGTTTATGACATCTGG 1200
1286 CTGCTTTGGAATCTTTTGTGATGCGATGAGATCTTTCAATTAAGCTGAATAACAA 1345
1201 CCATCCTGAGGAAACATCATGATGATGATCTTAATACCAAGCCAAATATAACTCAG 1260
1346 CTCACGCTATGCTCAAAATCTTGTGCTGATGCTGATGAGGTTTACTTGT 1405
1261 TCCCAAGTCTCTATATGCAACCTGGCTTGTGATCTTGCAATGGAATCTACCTGC 1320
1406 TCTTTGTTGGCATTTTGCATATATAAATACCGAGGCGAGTATCAGAATGCTTGTCT 1465
1321 TGTCTATGATGATGATGATATATCCATACCAAGCCAAATATCAACATATGCAATG 1380
1466 GGATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1525
1381 GGCAAACTGGGGCAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1526 CTGCTCTGCTACTGACCTACTGATCTTGGAGAGTTCTGCTGAT 1571
1441 CAGTCTACACTGTGACGACTATCCTTGGAAAGATGCAATACAT 1486

RESULT 4
US-08-482-855-1
Sequence 1, Application US/08482855
Patent No. 6121016

GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016en Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Stephen P. Williams,
ADDRESSEE: Ares-Serono, Inc.,
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Mase11 via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482, 855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670, 085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:

NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2179
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Testis
IMMEDIATE SOURCE:
LIBRARY: 19711 cDNA library, Clontech #HL1010b
CLONE: pF5HR11-11, pF5HR15-6
FEATURE:
NAME/KEY: protein coding region
LOCATION: 75 to 2159
US-08-482-855-1

Query Match 2.8%; Score 63.6; DB 3; Length 2179;
Best Local Similarity 47.3%; Pred. No. 3.1e-08;
Matches 192; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

1166 ACTGCTCTATGCTCCCATGTCGAATATGATGACCTTGAACGAGGCAATTTCTTAT 1225
1081 ACTATGCAATGAAGTGTATACGACCTGCTCCCAAGCCAGATGCAATCAACCAT 1140
1226 TTGAGGACCTTGGCTTACAAATATCTCAGAAATTTGTGGGTTTATGCTTATTA 1285
1141 GTGAAGATATCATGGGGTACAAACATCTCAGAGTCCATATGTTTATGACATCTGG 1200
1286 CTGCTTTGGAATCTTTTGTGATGCGATGAGATCTTTCAATTAAGCTGAATAACAA 1345
1201 CCATCCTGAGGAAACATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1346 CTCACGCTATGCTCAAAATCTTGTGCTGATGCTGATGAGGTTTACTTGT 1405
1261 TCCCAAGTCTCTATATGCAACCTGGCTTGTGATCTTGCAATGGAATCTACCTGC 1320
1406 TCTTTGTTGGCATTTTGCATATATAAATACCGAGGCGAGTATCAGAATGCTTGTCT 1465
1321 TGTCTATGATGATGATGATATATCCATACCAAGCCAAATATCAACATATGCAATG 1380
1466 GGATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1525
1381 GGCAAACTGGGGCAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1526 CTGCTCTGCTACTGACCTACTGATCTTGGAGAGTTCTGCTGAT 1571
1441 CAGTCTACACTGTGACGACTATCCTTGGAAAGATGCAATACAT 1486

RESULT 5
US-08-474-986-1
Sequence 1, Application US/08474986
Patent No. 6372711

GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6372711en Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Stephen P. Williams,
ADDRESSEE: Ares-Serono, Inc.,
STREET: Exchange Place, 37th floor

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Mass81 via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,986
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephen P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2179
TYPE: Nucleic acid
STRANDEDNESS: Double
MOLECULE TYPE: Linear
ORIGINAL SOURCE: CDNA to mRNA
ORGANISM: Homo sapiens
TISSUE TYPE: Testis
IMMEDIATE SOURCE:
LIBRARY: Igt11 CDNA library, Clontech #HL1010b
CLONE: PHFSHR11-11, PHFSHR15-6
FEATURE:
NAME/KEY: protein coding region
LOCATION: 75 to 2159
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-474-986-1
Query Match 2.8%; Score 63.6; DB 4; Length 2179;
Best Local Similarity 47.3%; Pred. No. 3.1e-08;
Matches 192; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

1166 ACTGCTCTATGCTCCCATGTCGGAATATGTAATGCCCTTGAGCGAGCGGATTTCTTCAT 1225
1081 ACTTATGCAATGAAAGTGTGACGTGACCTGCTCCCTTAAGCCAGATTCACCCAT 1140
1226 TTGAGGACCTCTTGAGCTAACATATCTCAGAAATATTTGTCTGGGTATAGCTTTCATTA 1285
1141 GTGAAGATATCAGGGGTACACATCTCTCAGAGTCCGATATGTTTATACAGATCTCG 1200
1286 CTGCTTTGGAATCTTTTGTTCATTTGGCATGAGATCTTTCATTAAAGTGAATAACAA 1345
1201 CCATCACTGGGAACATCATAGTGTAGTATCTTAACACAGCAATATATACTCAG 1260
1346 CTACGCTATGTCATCAAAATCCTTTGTGTCGTATGCTGATGCGGATTTACTCTT 1405
1261 TCCCCAGGTCCTTATGTGCACTGGCCTTGTGTGATCTCTGCAATGAAATCTACCTGC 1320
1406 TCTTTGTGGCATTTCGATATATAACGAGGCGAGTATCAGAAATATGCTTGTCTGT 1465
1321 TGCTCATTTGATGATGATATCATACCAAGACCAATATCAACAATATAGCCATTGACT 1380
1466 GGAATGAGAGCTGCAATGCGGCTCATGAGGATTCCTGGGCAATCTGTCCACGAAATCT 1525
1381 GGCAAACTGGGGCGAGGCTGTGATGCTGCTTTTTCATGCTTTTTCAGAGTGAAGCTGT 1440
1526 CTGTTCTGCTACTGACCTACTTGAATTTGAGAGATTTCTGGTCAAT 1571
1441 CAGCTCATCTCTGACAGCTATCAGCTTTGGAAGAAATGGAATACCAT 1486

RESULT 6
US-09-016-434-1209
Sequence 1209, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1209:
SEQUENCE CHARACTERISTICS:
LENGTH: 2393 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g182770
US-09-016-434-1209
Query Match 2.8%; Score 63.6; DB 4; Length 2393;
Best Local Similarity 47.3%; Pred. No. 3.3e-08;
Matches 192; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

1166 ACTGCTCTATGCTCCCATGTCGGAATATGTAATGCCCTTGAGCGAGCGGATTTCTTCAT 1225
1073 ACTTATGCAATGAAAGTGTGACGTGACCTGCTCCCTTAAGCCAGATTCACCCAT 1132
1226 TTGAGGACCTCTTGAGCTAACATATCTCAGAAATATTTGTCTGGGTATAGCTTTCATTA 1285
1133 GTGAAGATATCAGGGGTACACATCTCTCAGAGTCCGATATGTTTATACAGATCTCG 1192
1286 CTGCTTTGGAATCTTTTGTTCATTTGGCATGAGATCTTTCATTAAAGTGAATAACAA 1345
1193 CCATCACTGGGAACATCATAGTGTAGTATCTTAACACAGCAATATATACTCAG 1252
1346 CTACGCTATGTCATCAAAATCCTTTGTGTCGTATGCTGATGCGGATTTACTCTT 1405
1253 TCCCCAGGTCCTTATGTGCACTGGCCTTGTGTGATCTCTGCAATGAAATCTACCTGC 1312
1406 TCTTTGTGGCATTTCGATATATAACGAGGCGAGTATCAGAAATATGCTTGTCTGT 1465
1313 TGCTCATTTGATGATGATATCATACCAAGACCAATATCAACAATATAGCCATTGACT 1372

QY 1466 GGATGAGAGCGGTGAGTCCGCTCATGGGGTTCCTGGCCATGCTGTCCACCGAAGTCT 1525
DB 1373 GGCAAACTGGGGCAGCGGTGTGATGCTGCGCTTTTTCACGTGTCTTGCCAGTAGCTGT 1432
QY 1526 CTGTCTCTACTGACCTACTTGAATTTGGAGAAATTCCTGCTCAT 1571
DB 1433 CAGCTACACTGTACAGCTATACCTTGGAAGATGCAATACAT 1478

RESULT 7

US-09-016-434-1469
; Sequence 1469, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1469:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2413 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENEBANK
; CLONE: 9903759
; US-09-016-434-1469

Query Match 2.8%; Score 63; DB 4; Length 2413;

Best Local Similarity 47.6%; Pred. No. 5e-08;
Matches 186; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 1193 TATGTATGCGCTTGAAGGAGGCAATTTCTTATTTGAGAGCACTTGGCTAACAATATCC 1252
DB 1260 TGTGTACCCCAAGTCCGATGATTCACCGGTGGAAGACATATATGGGCTACAGTTCC 1319
QY 1253 TCAGAAATTTTGTCTGGGTATAGCTTTATTAACCTGCTTTGGAAATCTTTTGTGATTC 1312
DB 1330 TGAGAAATTTGTGTGTGCTTGTAGTCTGCTGCTCTCTGGGCAATGCTTTGTCTGTC 1379
QY 1313 GCATAGATCTTTCAATTAAGCTGAATAATACATCAAGCTATGTCATCAAAATCTTT 1372
DB 1380 TTATTTCTCTACACAGCCACTACAAATGAAAGTCTCCGCTTTCTCATGTGCAACTGG 1439

QY 1373 GTTGTGATATTCCTATATGAGGTGTTTACTTGTCTTTGTTGGATTTTCATATAAAT 1432
DB 1440 CCTTTGGAGATTTTGATATGGGAGATGATACCTGCTCTCATGCGCTGTATACCTTACA 1439
QY 1433 ACCGAGGCAATATCAGAAATATGCTTGTCTGTGATGAGAGAGCTGACGCGCTCA 1492
DB 1500 CTCACCTGTGATCTACAAACCATGCAATGCACTGGCAGACAGGCCCTGGGTGCAACGG 1559
QY 1493 TGGGTTCCCTGGCCATGCTGCCACGAAGTCTGTTGCTACTGACCTACTTGACTT 1552
DB 1560 CTGTTTCTTCACTGCTTTGCAAGCAGATGATGCTGTATAGCTGACGCTATACACC 1619
QY 1553 TGGAGAAATTCCTGTCATGTTCTTCCCTT 1583
DB 1620 TGAAGCGCTGTATGCCATACCTTCGCAAT 1650

RESULT 8

US-07-741-453A-62
; Sequence 62, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3710 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-07-741-453A-62

Query Match 2.8%; Score 63; DB 3; Length 3710;

Best Local Similarity 47.6%; Pred. No. 6.5e-08;
Matches 186; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 1193 TATGTATGCGCTTGAAGGAGGCAATTTCTTATTTGAGAGCACTTGGCTAACAATATCC 1252
DB 1522 TGTGTACCCCAAGTCCGATGATTCACCGGTGGAAGACATATATGGGCTACAGTTCC 1581
QY 1253 TCAGAAATTTTGTCTGGGTATAGCTTTATTAACCTGCTTTGGAAATCTTTTGTGATTC 1312

Db 1582 TGAGAAATGNGTGGTTCGTTAGTCTGTGCTCTCTGAGCAATGCTTGTCTGTC 1641
Qy 1313 GCATGAGATCTTTCAATTAAGCTGAATAATCAACGTATGTCATCAAAATCTTT 1372
Db 1642 TTATCTCTCACCACCACTACAACTGAACGTCCCGCTTCTCATGTGCAACCTG 1701
Qy 1373 GTTGTGCTGATTCCTGATGAGTGTGTTACTTGTCTTGTGTCATTTGATATTAAT 1432
Db 1702 CCTTGGGATTTCTGATGAGGATGATGCTGCTCTCATGCTGTGATGACCTTACA 1761
Qy 1433 ACCGAGGAGATACAGAAATATGCTTGTCTGTGATGAGAGCGTGACGCTGCTCA 1492
Db 1762 CTCACCTGAGTACTACCAACCATGCACTGAGAGAGCGCTGAGTGAACAG 1821
Qy 1493 TGGGGTTCCTGGCAGCTGTCCACGAGTCTGTCTGTCTGATGACCTTACTT 1552
Db 1822 CTGCTTCTTCACTGCTTTTTCAGAGGATTAAGTGTATGAGCTGACGCTATCACC 1881
Qy 1553 TGGAGAGTTCCTGTGATGCTTCCCTT 1583
Db 1882 TGAAGCGCTGTATGCACTACCTTGCCAT 1912

RESULT 9

US-07-741-453A-57
; Sequence 57, Application US/07741453A
; Patent No. 6228597

GENERAL INFORMATION:

APPLICANT: PARMENTIER, MARC
APPLICANT: LIBERT, FREDERIC
APPLICANT: DUMONT, JACQUES
APPLICANT: VASPART, GILBERT
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,453A
FILING DATE: 19911015
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714637 CUSH
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 4417 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-741-453A-57

Query Match

Best Local Similarity 47.6%; Score 63; DB 3; Length 4417;
Matches 186; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

Qy 1193 TATGATGCCCTTAAGCAGCGCATTTCTTCAATTGAGGACCTCTTGCTAACAATATCC 1252

Db 1252 TGTGTACTCTTAAGTCAGATGATTCACACCCCTGGAAGACATTAATGGGCTACAGTTCC 1311
Qy 1253 TCAGAAATATTTCTGCGGTATAGCTTCAATTAACCTGCTTGGAAATCTTTTTCATTG 1312
Db 1312 TGAGATTTGT 1371
Qy 1313 GCATGAGATCTTTCAATTAAGCTGAATAATCAACTCAGGCTATGTCATCAAAATCTTT 1372
Db 1372 TGTCTCTCTTACAGTCACTACATAATGATCTGTCCACGCTTCTCATGTGCAACTTGG 1431
Qy 1373 GTTGTGCTGATTCCTGATGAGTGTGTTACTTGTCTTGTGTCATTTGATATTAAT 1432
Db 1432 CCTTGCAGATTTCTGATGAGGATGATGCTGCTCTCATGCTCCGTAAGACCTTACA 1491
Qy 1433 ACCGAGGAGATACAGAAATATGCTTGTCTGTGATGAGAGCGTGACGCTGCTCA 1492
Db 1492 CTCATTTGAGTACTACCAACCATGCACTGAGAGAGCGCTGAGTGAACAG 1551
Qy 1493 TGGGGTTCCTGGCAGCTGTCCACGAGTCTGTCTGTCTGATGACCTTACTTACTT 1552
Db 1552 CTGCTTCTTCACTGCTTTTTCAGAGGATTAAGTGTATGAGCTGACGCTATCACC 1611
Qy 1553 TGGAGAGTTCCTGTGATGCTTCCCTT 1583
Db 1612 TGAAGCGCTGTATGCACTACCTTGCCAT 1642

RESULT 10

US-09-131-648-4
; Sequence 4, Application US/09131648
; Patent No. 6168920

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
FILE REFERENCE: PF-0576 US
CURRENT APPLICATION NUMBER: US/09/131,648
CURRENT FILING DATE: 1998-08-10
SOFTWARE: PERL Program
SEQ ID NO 4
NUMBER OF SEQ ID NOS: 5
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 2687731
US-09-131-648-4

Query Match

Best Local Similarity 45.0%; Score 60.6; DB 3; Length 2290;
Matches 228; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

Qy 425 TGTCTCTTAAGAAAACAAATCCACAGTCTTCAGATTAAGTTTTCATCAATAACAA 484
Db 424 TGTACTAGAGAAAACAACTTACGAACTGCTGAAAATGTGTGCAACTGAGCA 483
Qy 485 AACTTAAGATATTTCTTACGATTAATGCTTGAACATATGACATATGCAAGAAACATTT 544
Db 484 ACTTACAAGAACTCTATTAATTAACAACTTGCTTTCTCAATTTCACTGAGCCCTTA 543
Qy 545 TTGATTAATGATATGCAAAATTAATATCTCAACACAGACTGCATACAACTCGAG 604
Db 544 TTGCTTCACTATCTTTCTTCACTTCAATTTCAATTTGAGATGAGATGACAA 603
Qy 605 CTGAATATTTCAAGACTTACATCACTGCTTAACTTGAATTTGATGACATTCATTA 664
Db 604 GTAAAGGTTTGAATGCTTCCAAATCTAGAGATTCGATGATGGGAAAATTCATTA 663
Qy 665 CCAGAATTTACAGCGCTGTATTAGCGGATTAATTCCTTGTTCCTGTCTATAGTTA 724

Db 664 TCAGATATGAAGACATGAACTTTAAGCCTCTTATCAATCTTCGACGCGCTGTATAGCTG 723
Qy 725 ATAACTACTTGAAGAGCTCTTCCCAAGAGATGTGCGCAATGCGCTCAACTCAACTGAG 784
Db 724 GTATTAACCTCAGCAAAATACAGATTAACGCTTGTTGAGACTGAGAAAATTAGAAACA 783
Qy 785 TGAATTGGAAGGCAATGAAATTAAGTATCTCAAAATCTCAAGTTCTGTGCGCAT 844
Db 784 TCTCTTTTACGATTAACAGGCTTATTAAGTACCCCATGTTGCTCTCAAAAAGTTGTA 843
Qy 845 CGCTCAGAGTGTGTTCTGCTTGAAGAAATCAAAATGTTGTTTCCAGAGAAAGCATTTT 904
Db 844 ATCTCAAAATTTTGGATCTAAATTAATAAAATCTTATTAATGAATACGAAGGGGTGATTTTA 903
Qy 905 CTTCAATTAATAATTTAGAGAACTGG 931
Db 904 GCAATATGCTACCTTAATAAGAGTTGG 930

RESULT 11
US-09-620-312D-350

; Sequence 350, Application US/09620312D

; Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungting
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghaast
APPLICANT: Drmanac, Radote T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_Fl_genes Version 1.0
SEQ ID NO 350
LENGTH: 2461
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (232)..(2358)
US-09-620-312D-350

Query Match 2.7%; Score 60.6; DB 4; Length 2461;
Best Local Similarity 45.0%; Pred. No. 2.5e-07;
Matches 228; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

Qy 425 TGTCTTTAGAGAAACAAATCCACAGCTTCCAGATTAAGTTTTCATCAATACACA 484
Db 593 TGTACTTAGAGAAACAACTACTGAACTGCTGAAAATGCTGTCGAACTGAGCA 652
Qy 485 AACTTAAAGATATTTCTTCAGCATTAATGATTAAGACACATATCCAGAAAGCATTTT 544
Db 653 ACTTCAAGAACTATATTAATCACAACTTGTCTTCAATTTCACTGAGAGCTTTA 712

Qy 545 TTGATTATGTATCTGCAATATTAATATCAACCAACTGACATCAACCTCAGAC 604
Db 713 TTGGCTACATATATCTTCTGATCTCATCTCAATTAATAATGATGATGATCAACA 772
Qy 605 CTGAATATTTAAAGACTTACATCACTAGCTAATCTGGCTTAATCTAGATGACATCAATTA 664
Db 773 GTAAGTGGTTGAAGCTCTTCCAAATCTAGAGATTCGATGATGGGAAAATCAATTA 832
Qy 665 CCAGAAATTTACAGAGGCTGTTTACGGGATTAATATTCCTGTTTCCCTGATGTTA 724
Db 833 TCAGAAATCAAGACATGAACTTTAAGCCTTTATCAATCTTCGACGCTGTTATAGCTG 892
Qy 725 ATAACTACTTGAAGAGCTCTTCCCAAGAGATGTGCGCAATGCGCTCAACTCAACTGAG 784
Db 893 GTATTAACCTCAGCAAAATACAGATTAACGCTTGTTGAGACTGAGAAAATTAGAAACA 952
Qy 785 TGAATTGGAAGGCAATGAAATTAAGTATCTCAAAATCTCAAGTTCTGTGCGCAT 844
Db 953 TCTCTTTTACGATTAACAGGCTTATTAAGTACCCCATGTTGCTCTTCAAAAAGTTGTA 1012
Qy 845 CGCTCAGAGTGTGTTCTGCTTGAAGAAATCAAAATGTTGTTCCAGAGAAAGCATTTT 904
Db 1013 ATCTCAAAATTTTGGATCTAAATTAATAAAATCTTATTAATGAATACGAAGGGGTGATTTTA 1072
Qy 905 CTTCAATTAATAATTTAGAGAACTGG 931
Db 1073 GCAATATGCTACCTTAATAAGAGTTGG 1099

RESULT 12
US-08-985-335-4

; Sequence 4, Application US/08985335

; Patent No. 6080847

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
TITLE OF INVENTION: PROLIFERATION
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,335
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0421 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LATRUT02
CLONE: 1352286
US-08-985-335-4

Query Match 2.6%; Score 57.8; DB 3; Length 2082;
Best Local Similarity 51.8%; Pred. No. 1.5e-06;
Matches 131; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 407 CTACATGTGACATTAAGTCTCTTAAAGAAAAACAATCCAGCTCTTCAGATTAAG 466
DB 505 CTCAGAGTTTGCGACATGCTGCTAGCAAGAAACAGATCCGACATTAAGAACACA 564
QY 467 TTTTCATCAATACACAAAATTAAAGATTTTCTGACATTAATGATTAAGACA 526
DB 565 TGTTCCTCAAGTTTAAAGAAAGCTGAATCTGACAGAAAGAGATCTTAAAA 624
QY 527 TATCCAGAAAGATTTTGGATTATGTAATCTGCAAAATATTATCTCAACCAACT 586
DB 625 TTGAGGTGAGGCTTCTTTGGTTTAAACAACCTCACCCTCTTACTGACACACC 684
QY 587 GCATCAACAACCTCAGACCTGGAATATTCAAAGCTTACATCAGCTAATGCTAATTC 646
DB 685 AGATCAAGCTTGACGAGAAAGTGTTCATTTACACCTCTTGAGCTACCTGCTC 744
QY 647 TAGATGACATCC 659
DB 745 TTTATGACAACC 757

RESULT 13

US-09-410-372-4

Sequence 4, Application US/09410372
Patent No. 6281334
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Lai, Preeti
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/410,372
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/985,335
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0421 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LATRUT02
CLONE: 1352286
US-09-410-372-4

Query Match 2.6%; Score 57.8; DB 3; Length 2082;
Best Local Similarity 51.8%; Pred. No. 1.5e-06;
Matches 131; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 407 CTACATGTGACATTAAGTCTCTTAAAGAAAAACAATCCAGCTCTTCAGATTAAG 466
DB 505 CTCAGAGTTTGCGACATGCTGCTAGCAAGAAACAGATCCGACATTAAGAACACA 564
QY 467 TTTTCATCAATACACAAAATTAAAGATTTTCTGACATTAATGATTAAGACA 526
DB 565 TGTTCCTCAAGTTTAAAGAAAGCTGAATCTGACAGAAAGAGATCTTAAAA 624
QY 527 TATCCAGAAAGATTTTGGATTATGTAATCTGCAAAATATTATCTCAACCAACT 586
DB 625 TTGAGGTGAGGCTTCTTTGGTTTAAACAACCTCACCCTCTTACTGACACACC 684
QY 587 GCATCAACAACCTCAGACCTGGAATATTCAAAGCTTACATCAGCTAATGCTAATTC 646
DB 685 AGATCAAGCTTGACGAGAAAGTGTTCATTTACACCTCTTGAGCTACCTGCTC 744
QY 647 TAGATGACATCC 659
DB 745 TTTATGACAACC 757

RESULT 14

US-07-757-342D-9

Sequence 9, Application US/07757342D
Patent No. 6218509
GENERAL INFORMATION:
APPLICANT: IGARASHI, Masao
APPLICANT: MINEGISHI, Takashi
APPLICANT: NAKAMURA, Kazuo
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRB UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2022 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

: MOLECULE TYPE: cDNA to mRNA
:
: FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 1..2022
:   SEQUENCE DESCRIPTION: SEQ ID NO: 9
US-07-757-342D-9

```

Query Match	2.5%	Score 57.6;	DB 3;	Length 2022;
Best Local Similarity	44.8%;	Pred. No. 1.7e-06;		
Matches 222; Conservative	0;	Mismatches 274;	Indels 0;	Gaps 0;

Qy	1081	GAAAGGATAGAGATCTCCAAATATAAACAAGAAAGTTCAACCATAAGAAATCTTCT	1140
Db	838	GAAAGCAAGTATAGGAAAGTGAAGTAAACAAACATTATCTTCATGCTGCTGAGAGT	897
Qy	1141	CACATTTATTTCAAAAATTTCGATAGTCTGTATGCTCCCATATGCCAATATGATATG	1200
Db	898	GAACCTGATGGCTGGGACTATGAATATGTTCTCTTACCCAGAACACCCCGATGTCT	957
Qy	1201	CCCTTGAACGACGGCATTTCTTCATTGAGGACCTCTTGCGCTAACATATCCTGAAAT	1266
Db	958	CTGACACCAAGATGCTTTAATCCCTGTAGAACATTATAGGGTATGACCTTCTTAGGGTC	1017
Qy	1261	TTTGTCTGGGTTATAGCTTTCATTACCTGCTTTGGAAACTTTTGTCAATGGACATGAGA	1326
Db	1018	CTGATTTGGCTGATTAATATCTTACCAATCATGGGAAACATGACGTGCTTTTGTCTC	1077
Qy	1321	TCTTTCAATAAAGCTGAATAATCAACTCAAGTATGTCATGAATCCTTGTGTGCT	1380
Db	1078	CTGACAAAGCTGTACAAACTTACAGTGCCTCGTTTCTATATGCAATCTCTCTTGA	1133
Qy	1381	GATTGCTGATGGGTGTACTTGTCTTGTGTGGCATTTTCGATATAAATATACGAGGG	1440
Db	1138	GACTTTTGCAATGGGGCTCTATCTGTGCTCATAGCTCAGTTGATTTCCAAACCAAGGGC	1197
Qy	1441	CAGTATCAGAAGTATAGCCTGTGCTGTGATGAGAGACGTGCAGTATGGCCCTCATAGGGTTC	1507
Db	1198	CAGTACTATTAACCATGACCATAGACTGGCAGACAGAGAGTGGGTGACGACTGCTGCTTT	1257
Qy	1501	CTGGCCATGCTGTCCACCGAAGTCTCTGTTCTGTCAATAGCACTTACCTTGACATTTGGAGAAG	1566
Db	1258	TTCACTGATATGGCAAGTGAATCTTCTGTCTACACCTCAGCGTATCACTTATAAAGA	1311
Qy	1561	TTCTCTGTCAATGTCT	1576
Db	1318	TGGCACACATCACT	1333

RESULT 15
 US-07-757-342D-1
 ; Sequence 1, Application US/07757342D
 ; Patent No. 6218509
 ; GENERAL INFORMATION:
 APPLICANT: IGARASHI, Masao
 MINEGISHI, Takashi
 NAKAMURA, Kazuo
 TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSSE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
 CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/757.342D

```

1      FILING DATE: 10-Sep-1991
2      CLASSIFICATION: <Unknown>
3      ATTORNEY/AGENT INFORMATION:
4      NAME: BUCKLEY, Linda M.
5      REGISTRATION NUMBER: 31003
6      REFERENCE/DOCKET NUMBER: 41226
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE: (617)523-3400
9      TELEFAX: (617)523-6440
10     TELEX: 200291 STRE UR
11
12     INFORMATION FOR SEQ ID NO: 1:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 2987 base pairs
15     TYPE: nucleic acid
16     STRANDEDNESS: double
17     TOPOLOGY: linear
18     MOLECULE TYPE: cDNA to mRNA
19     FEATURE:
20     NAME/KEY: CDS
21     LOCATION: 1..2097
22     SEQUENCE DESCRIPTION: SEQ ID NO: 1
23
24     US-07-757-342D-1

```

Query Match	2.5%	Score 57.6	DB 3	Length 2987
Best Local Similarity	44.8%	Pred. No. 2	1e-06	
Matches	222	Conservative	0	Mismatches 274; Indels 0; Gaps 0;
QY	1081	GAAAGATAGAGATTCCAAATATTAACAACAAGATGTTTCAACCCATGAAAGATCTTTC	1140	
DB	913	GAAAGCACAGTAAAGAAAGTAGTAACAAACACTTATTTCTTCACATGCTTGAGAGT	972	
QY	1141	CACATTTATTTCAAAAACCTTTCGATCTGCTCCTATAGTCTCCCAATGTCGAATATGATG	1200	
DB	973	GAACTGAGTGGCTGGGACTATGAATATGTTCTTGCTTACCCAAAGACACCCCGATGTC	1032	
QY	1201	CCCTTGACGACGGGCAATTTCTTCATTTGAGGACCTCTTGCTAACAATATCCTCAGAA	1260	
DB	1033	CCTGAACCAAGATGCTTTTAATCCCTGGBAAGACTATATGGCTATGACTTCTTATGGGTC	1092	
QY	1261	TTTGTCTGGGTTATAGCTTCATTACCTGCTTGGAAATCTTTTGTCAATGGCAGTGA	1320	
DB	1093	CTGATTTGGCTGATTAATATCTAGCCATCATGGGAAACAGACGTGTTCTTTTGTGTC	1152	
QY	1321	TCTTTCAATTAAGCTGAAAAATACAATCAGCTATGTCATCAAAATCTTTGTTGTGCT	1380	
DB	1153	CTGACAAATCGTTCAAACTACAGTGCCTCGTTTTCTCAGTGCAAATCTCTCTTGCA	1212	
QY	1381	GATTCCTGATGGGGGNGTTACTTCTTCTTGGTGGCAATTTGCATPATAAATACCGAGG	1440	
DB	1213	GACTTTTGCAATGGGGCTTATCTGCTGCTCATAGCTCTGATGTAATCCCAACCAAGGCG	1272	
QY	1441	CAGATACGAAGTATGCCCTTGCTGTGATGAGAGCGTGCAGTCCCGCTCATGGGGTTC	1500	
DB	1273	CAGTACATTAACATGCCATAGACTGGGAGACAGGAGTGGGTGAGACATGCTGGCTT	1332	
QY	1501	CTGGCGATGCTGTCCACCGAAGTCTGTGTTGTGCTACTGACTTACTTGGAGAG	1560	
DB	1333	TTCACTGATTTCCGAAGTGAATCTTGTCTACACCCCTCACCGTCATCACTTAGAAGA	1392	
QY	1561	TTCCGTGATTTGTCT	1576	
DB	1393	TGGCACCAATACCT	1408	

Search completed: November 28, 2003, 19:19:49
Job time : 158 secs